233/27.8.70

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
BORIS SEMEONOFF

WITH THE ASSISTANCE OF
C. BURT

AIDED IN THE SELECTION OF PAPERS BY
F. C. BARTLETT
E. A. BOTT
R. S. CREED
J. DREVER
G. C. DREW
H. J. EYSENCK
L. S. HEARNSHAW
M. G. KENDALL
O. A. OESER
T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS

ASSISTANT EDITOR: IAN M. L. HUNTER

VOLUME 50, 1959

BUREAU OF EDNL. & PSYL. RESEARCH : : LIBRARY

CAMBRIDGE
AT THE UNIVERSITY PRESS
1959

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
Bentley House, 200 *Euston Road, London, N.W.* 1
American Branch: 32 *East 57th Street, New York* 22, *N.Y.*
Agent in Canada: The Macmillan Company of Canada Limited

Bureau Edni. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated 27. 8. 70
Accs No g 233

CONTENTS OF VOL. 50

Part 1. February 1959

PAGE

Explanatory powers of the fractional antedating response mechanism. By Benbow F. Ritchie 1

Quantitative judgement in social perception. By H. Tajfel 16

Attention to visual displays during a vigilance task. II. Maintaining the level of vigilance. By C. H. Baker 30

Increase in perceptual sensitivity as a function of learning in the test situation. By J. D. Keehn 37

Figural after-effects, retinal size and apparent size. By Peter McEwen . 41

Eysenck's theory of anxiety and hysteria—a methodological critique. By Vernon Hamilton 48

Anxiety and hysteria—a reply to Vernon Hamilton. By H. J. Eysenck . 64

Note on Burt's 'The Distribution of Intelligence'. By J. C. Raven . . 70

The 'Phenomenological Fallacy'—a reply to J. R. Smythies. By U. T. Place 72

Publications recently received 74

Other publications received 88

Part 2. May 1959

An objective investigation of the dynamics of semantic systems. By A. R. Luria and O. S. Vinogradova 89

The stroboscopic patterns. I. The dark phase. By J. R. Smythies . . . 106

The Tolman and Honzik insight situation. By W. S. Anthony 117

The effect of a warning signal on unrestricted avoidance behaviour. By J. D. Keehn 125

Effects of age and experience on the establishment of internal inhibition in finches. By M. A. Vince 136

Effects of television on the interests and initiative of adult viewers in Greater London. By William A. Belson 145

Development of the perception of social differences in children from 6 to 10. By G. Jahoda 159

The differentiation between normal and various neurotic groups on the Maudsley personality inventory. By H. J. Eysenck 176

Publications recently received 178

Other publications received 191

PART 3. AUGUST 1959

PAGE
The solving of five-letter anagram problems. By IAN M. L. HUNTER . . . 193
Remembering advance cues during searching. By N. H. MACKWORTH and J. F. MACKWORTH . . . 207
Hue differences and brightness differences as determinants of figural after-effects. By R. H. DAY . . . 223
Inter-sensory perceptual set. By GEORGE A. TALLAND . . . 231
Objective-test assessment of the primary personality dimensions in middle childhood. By RAYMOND B. CATTELL and RICHARD W. COAN . . . 235
Positive and negative information in matching problems. By MARGARET DONALDSON . . . 253
An interpretation of the 'displacement' phenomenon. By DALBIR BINDRA . . . 263
The paired use of projective stimuli. By JOHN LIGGETT . . . 269
Theories of anxiety and hysteria—a rejoinder to Hans Eysenck. By VERNON HAMILTON . . . 276
Corrigendum . . . 280
Publications recently received . . . 281

PART 4. NOVEMBER 1959

Effect of prejudice on reasoning. By ROBERT H. THOULESS . . . 289
The anchoring effects of value in a scale of judgements. By H. TAJFEL . . . 294
The stroboscopic patterns. II. The phenomenology of the bright phase and after-images. By J. R. SMYTHIES . . . 305
Extraversion-Introversion and improvement in an auditory vigilance task. By PAUL BAKAN . . . 325
Lability in schizophrenia. By N. O'CONNOR and J. P. DAS . . . 333
The blocking effect of benactyzine hydrochloride on a behaviour disturbance induced with lysergic acid diethylamide. By PETER H. GLOW . . . 338
Errors of immediate memory. By R. CONRAD . . . 349
Serial position effects in nonsense syllable learning as a function of interlist rest pauses. By H. J. EYSENCK . . . 360
Inhelder and Piaget's *The Growth of Logical Thinking*. I. A psychologist's viewpoint. By JEROME S. BRUNER . . . 363
Publications recently received . . . 371

INDEX OF AUTHORS

PAGE
Anthony, W. S. 117
Bakan, P. 325
Baker, C. H. 30
Belson, W. A. 145
Bindra, D. 263
Bruner, J. S. 363
Cattell, R. B. 235
Coan, R. W. 235
Conrad, R. 349
Das, J. P. 333
Day, R. H. 223
Donaldson, M. 253
Eysenck, H. J. 64, 176, 360
Glow, P. H. 338
Hamilton, V. 48, 276
Hunter, I. M. L. 193
Jahoda, G. 159
Keehn, J. D. 37, 125
Liggett, J. 269
Luria, A. R. 89
Mackworth, J. F. 207
Mackworth, N. H. 207
McEwen, P. 41
O'Connor, N. 333
Place, U. T. 72
Raven, J. C. 70
Ritchie, B. F. 1
Smythies, J. R. 106, 305
Tajfel, H. 16, 294
Talland, G. A. 231
Thouless, R. H. 289
Vince, M. A. 136
Vinogradova, O. S. 89

INDEX OF PUBLICATIONS REVIEWED

Ahrenfeldt, R. H. *Psychiatry in the British Army in the Second World War* 282
Ainsworth, L. H. & M. D. *Measuring Security in Personal Adjustment* 285
Beardslee, D. C. & Wertheimer, M. *Readings in Perception* 83
Brachmachari, S. *Social and Moral Attitudes* 378
Brain, Sir R. *The Nature of Experience* 178
Broadbent, D. E. *Perception and Communication* 78
Burt, Sir C. *A Psychological Study of Typography* 281
Carthy, J. D. *An Introduction to the Behaviour of Invertebrates* 288
Cazacu, T. S. *Relatiile dintre Gandire si Limbaj in Ontogeneza* 185
Chapanis, A. *Research Techniques in Human Engineering* 373
Chesser, E. *An Outline in Human Relationships* 378
Clarke, A. D. B. & A. M. *Mental Deficiency: The Changing Outlook* 375
Clostermann, G. *Die Prozent-Quotient-Methode* 286
Crook, G. H. & Heinstein, M. *The Older Worker in Industry* 284
Deese, J. *The Psychology of Learning* 374
Diehl, C. F. *Stuttering* 287
Educational Research, Vol. I, No. 1 183
Emery, F. E. & Oeser, O. A. *Information, Decision and Action: A Study of the Psychological Determinants of Changes in Farming Techniques* 81
English, A. C. & H. B. *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms* 86
Farnsworth, P. R. *The Social Psychology of Music* 84
Fleming, C. M. *Teaching: a Psychological Analysis* 183
Frankenstein, C. *Die Äusserlichkeit des Lebensstils (ein Beitrag zur Pathologie der Ichentwicklung)* 384
Franklin, E. *Tonality as a basis for the study of Musical Talent* 84
French, T. M. *The Integration of Behavior.* Vol. III: *The Reintegrative Process in a Psychoanalytic Treatment* 285
Garattini, S. & Ghetti, V. *Psychotropic Drugs* 85
Garrett, H. E. *Statistics of Psychology and Education* 182
Goldberg, E. M. *Family Influences and Psychosomatic Illness: An Inquiry into the Social and Psychological Background of Duodenal Ulcer* 286
Graefe, O. *Structuren der Rundfunkprogramm-Auswahl bei Hausfrauen* 84
Gregory, C. C. L. & Kohsen, A. *The O-Structure: An Introduction to Psychophysical Cosmology* 380
Hakkinen, S. *Traffic Accidents and Driver Characteristics* 182
Hansen, K. H. *After-Effects in the Behaviour of Mice* 180
Harlow, H. F. & Woolsey, C. N. *Biological and Biochemical Bases of Behavior* 375
Hebb, D. O. *A Textbook of Psychology* 186
Heider, F. *The Psychology of Interpersonal Relations* 82
Henry, H. *Motivation Research* 383
Hilgard, E. R. *Theories of Learning* 374
Himmelweit, H. T., Oppenheim, A. N. & Vince, P. *Television and the Child: An Empirical Study of the Effect of Television on the Young* 281
Hutchinson Medical Publications. *The Nature of Stress Disorder* 384
Jacobson, E. H. *International Directory of Psychologists Exclusive of the U.S.A.* 288

Janis, I. L. *Psychological Stress: Psychoanalytic and Behavioral Studies of Surgical Patients* 283
Kerr, M. *The People of Ship Street* 184
Kissen, D. M. *Emotional Factors in Pulmonary Tuberculosis* 87
Leighton, A. H., Clausen, J. A. & Wilson, R. N. *Explorations in Social Psychiatry* 184
MacPherson, J. S. *Eleven-year-olds Grow Up* 189
Mautner, H. *Mental Retardation: Its Care, Treatment and Physiological Base* 376
McIlwain, H. *Chemotherapy and the Central Nervous System* 85
McKellar, P. *Imagination and Thinking: a Psychological Analysis* 74
Melhuish, G. *The Paradoxical Universe* 382
Menninger, K. *Theory of Psychoanalytic Technique* 86
Meyer, L. B. *Emotion and Meaning in Music* 84
Peters, R. S. *The Concept of Motivation* 187
Raven, J. C. *Guide to using the Coloured Progressive Matrices, Sets A, Ab, B* 186
Raven, J. C. *Guide to using the Mill Hill Vocabulary Scale with the Progressive Matrices scales* 186
Reed, C. F., Alexander, I. E. & Tomkins, S. S. *Psychopathology: a Source Book* 178
Revista de Psihologie 185
Rice, A. K. *Productivity and Social Organisation: The Ahmedabad Experiment* 190
Robertson, J. R. *Young Children in Hospital* 182
Russell, W. R. *Brain—Memory—Learning: a Neurologist's View* 178
Schiller, C. H. *Instinctive Behavior: The Development of a Modern Concept* 287
Scott, J. P. *Aggression* 180
Semeonoff, B. & Trist, E. *Diagnostic Performance Tests: A Manual for Use with Adults* 80
Senders, V. L. *Measurement and Statistics. A Basic Text Emphasizing Behavioral Science Applications* 188
Skinner, B. F. *Verbal Behavior* 371
Sprott, W. J. H. *Human Groups* 179
Stone, A. A. & Onque, G. C. *Longitudinal Studies of Child Personality: Abstracts with Index* 377
Stott, D. H. & Sykes, E. G. *The Bristol Social Adjustment Guides* 379
Tagiuri, R. & Petrullo, L. *Person Perception and Interpersonal Behavior* 82
Thomson, R. *The Psychology of Thinking* 382
Thurstone, L. L. *The Measurement of Value* 377
Wall, W. D. *Educational Research* 183
Welford, A. T. *Ageing and Human Skill* 77
Wilmer, H. A. *Social Psychiatry in Action: a Therapeutic Community* 187
Winnicott, D. W. and others. *Psycho-Analysis and Contemporary Thought* 179
Wolfgang de Boor. *Pharmakopsychologie und Psychopathologie* 85
Wolstenholme, G. E. W. & O'Connor, C. M. *Ciba Foundation Symposium on the Neurological Basis of Behaviour* 79
Ziler, H. *Der Mann-Zeichen-Test* 184

VOLUME 50 FEBRUARY 1959 PART 1

EXPLANATORY POWERS OF THE FRACTIONAL ANTEDATING RESPONSE MECHANISM

By BENBOW F. RITCHIE

Psychology Department, University of California, Berkeley

The fractional antedating response mechanism is used by Spence to explain the fact that rats can anticipate the consequences of their responses and can make choices based on these anticipations. The present paper examines the logic of this explanation, shows the kinds of assumptions required by it, and derives an experimental test of this explanation.

This derivation assumes that fractional eating reactions in black and in white boxes are different reactions. While this assumption is not required to explain Kendler's experiment or Spence's paradigm, it is required to explain the latent learning and reasoning phenomena in many other experiments.

As a consequence, the negative results of the experiment reported in this paper suggest either that the fractional reaction explanation is false, or that it is trivial in the sense that it is appropriate only to those latent learning phenomena where the alternative consequences are distinguished by different consummatory responses.

I. INTRODUCTION

This paper has two purposes. (1) To provide a *complete* r_G explanation of one experiment (Kendler, 1946). (2) To test this explanation by another experiment. Until we have succeeded in the first purpose, we cannot hope to succeed in the second. Such an exercise in argument and experiment may give us a better understanding of the explanatory powers of the r_G.

Kendler (1946) trained rats in a simple T maze with food at the end of one arm and water at the end of the other. Half of the maze, including the goal box, the maze arm, and that side of the starting stem, was painted black; the other half was left unpainted. After preliminary training, in which the animals were allowed to explore the maze (no food or water present), the animals were made 21 hours hungry and 21 hours thirsty.

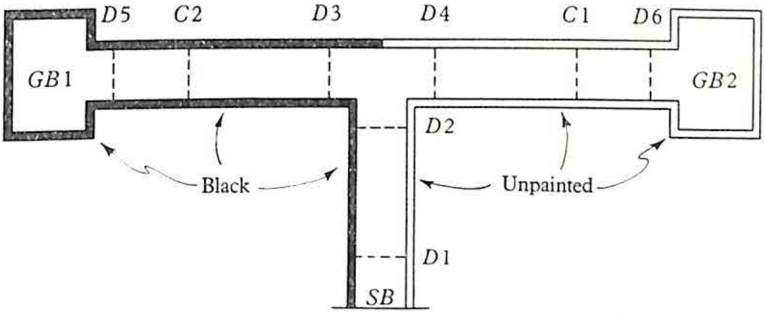


Fig. 1. Kendler's maze. *SB* indicates starting box; *D*1, *D*2, etc., doors; *C*1 and *C*2, curtains; and *GB*1 and *GB*2 goal boxes.

The training series consisted of four daily trials for 7 days, with food at the end of one arm and water at the end of the other. The first run of each day was a free-choice trial. On the second trial, the animal was forced to the side opposed to his first choice. The third trial was another free choice, and on the fourth trial he was forced to the side opposed to that choice. Thus, on each day each animal had two runs in each alley. If there was an inequity in the reinforcing value of water and food, it was not enough to produce an alley preference.

The animals were then tested either hungry or thirsty to discover whether they would choose the alley appropriate to the single drive. Each animal was tested once on four successive days either *HTTH* or *THHT*. The results (see Fig. 2) show that the animals chose the alley leading to the appropriate incentive.

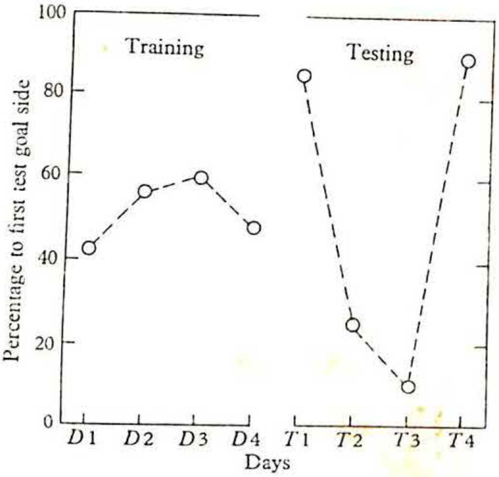


Fig. 2. Percentage of animals during training and testing which chose the maze arm appropriate on test trials $T1$ and $T4$.

II. SPENCE'S r_G EXPLANATION

We commonly assume that the anticipation of future consequences can somehow control our present behaviour. The Kendler experiment shows that such control is also important in the behaviour of rats. The psychological problem posed by this fact is a crucial one. If we are to understand the fact, we must conceive, first, of some mechanism which represents such future consequences, and secondly, of some way in which this mechanism controls the responses leading to different consequences. Spence believes that the fractional antedating response mechanism can explain this fact and has elaborated an argument to show how this mechanism represents the consequences and how it controls behaviour. The Kendler experiment is thus important for two reasons: (1) it illustrates a pure case of purposive behaviour, and (2) it exemplifies the explanatory power of the fractional antedating response mechanism.

For the sake of simplicity let us begin by considering a hungry rat trained to run a straight alley to a goal box containing food. When the rat is in the goal box, he makes RG (the goal response of eating the food). Such a goal response contains a fractional component r_G which can occur when no SG (food) is present. When this r_G occurs in the goal box, it is reinforced by $SG \to RG$, and becomes conditioned to traces from cues in the alley leading to the goal box. As a consequence of this trace-conditioning the cues at the entrance of the alley 'will evoke, through generalization, such components of this goal response (r_G) as can occur without the actual presence of the goal object' (Spence, 1951, p. 272). In this sense, the r_G antedates the RG to which it belongs; in this sense, and only in this sense, does the r_G 'represent the SG or the RG'; and in this sense, and only in this sense, does the r_G 'anticipate the consequences of traversing the alley'.

'The interoceptive stimulus cue (s_G) produced by this classical conditioned response will itself become conditioned to the instrumental locomotor response and thus become a determiner of its excitatory strength' (Spence, 1951, pp. 272–3). In this way the r_G provides a new kind of stimulus for the initial part of a response sequence. What is new about s_G is that the characteristics of s_G are determined by the characteristics of the SG which evokes RG at the end of the sequence. This, then, is the mechanism which represents the consequences of running down the alley.

Now, how does this stimulus influence the response of running down the alley? 'Presumably it does this in two ways: (1) the s_G acquires habit strength (s_GH_R) to the instrumental locomotor response, and (2) through the intensity of its trace, which presumably increases with increased strength of r_G as the latter becomes more strongly conditioned, s_G determines the non-associative factor, K. Recalling that

$${}_SE_R = f(K \times {}_SH_R),$$

it will be seen that this mechanism determines the excitatory strength of the instrumental response in two ways. That is, it contributes a component to ${}_SH_R$ and it determines the value of K' (Spence, 1951, p. 273).

We turn now to examine how Spence employs this hypothesis in his analysis of the Kendler experiment. Since in this case both choice responses were equally reinforced during training, Spence assumes 'that their habit strengths (${}_SH_R$) were strengthened to the same degree during training. How then', he asks, 'is the differential response on the test trials to be explained?' (1951, p. 278).

Spence assumes that:

> the cues at the entrance to the path leading to water through conditioning in the training series come to elicit the fractional anticipatory drinking act (r_W) and, in turn, the proprioceptive cue (s_W) from this act becomes conditioned to response of approaching and locomoting in this alley. Similarly, the fractional anticipatory eating response (r_F) becomes conditioned to the other alley entrance, and the proprioceptive cue (s_F) from it becomes conditioned to the behaviour of entering this alley. During the test series, when only one need is operative, the anticipatory act related to the goal object for which the subject is motivated, will, because of the greater strength of the particular drive stimulus (see the principle of stimulus dynamism) be stronger than the other, and hence will produce the stronger proprioceptive cue. For example, if the subject is thirsty, s_W will be stronger (more intense) than s_F. The intensity of these conditioned proprioceptive cues (s_W, s_F) determines, it will be recalled, the non-associative factor, K, and hence the excitatory strengths of the responses of entering and continuing down the paths to which they are conditioned. The greater intensity or strength of s_W when thirsty and not hungry leads to the implication that the excitatory potential of the response to the alley leading to water will be stronger than that of the response of entering the food alley. In other words, in the test series, the relative excitatory strengths of the competing responses will be a function of the relative strengths of the two fractional anticipatory goal responses and their proprioceptive cues (factor K). These latter will, in turn, be dependent upon which need is present and which is absent (1951, pp. 278–9).

This explanation is remarkable in several ways. It recognizes and deals with *both* aspects of the problem, the description of a 'representational' mechanism, and a description of how this mechanism controls behaviour. There is nothing in the argument which is inconsistent with the behaviourist premises that the elements of behaviour are S's and R's, and that the only kinds of changes that occur in the structure of behaviour are changes in the associative connexions between S's and R's. The only assumption which Spence has added to the basic conditioning principles is the assertion that the incentive motivation, K, is determined by the intensity of s_G. Finally, Spence's argument reveals the explanatory power of the r_G more forcefully than any other treatment of this concept in the literature primarily because the argument is more fully developed than any other employing the r_G.

But despite the excellence of Spence's argument it is not complete, and there are at least three points which need to be settled before we have a complete explanation of the Kendler experiment.

(1) *What, for example, prevents r_F from becoming conditioned to the cues in the water alley?* Clearly Spence's argument depends upon believing that r_F is not as strongly conditioned to the cues in the water alley as it is to the cues in the food alley. But this argument contains no proof that r_F does not also become conditioned to the water alley. Since the rats are trained both hungry and thirsty, and since every response to the cues in the water alley is reinforced, it seems that the r_F's evoked by the compound drive stimulus must, *from the first trial*, become conditioned to the cues in *both* alleys. The trace-conditioned r_G's, by hypothesis, can occur only after the habits reach threshold strength. This entails that the conditioned r_G's based upon r_G's evoked by the compound drive stimulus must have an irreversible advantage over those conditioned r_G's based upon trace conditioning. For this reason the Spence argument needs some proof that r_F is more strongly conditioned to the food alley cues.

(2) Assuming that r_F is conditioned solely to the cues in the food alley, another question arises. *What prevents s_F from becoming conditioned to entering the water alley?* Clearly Spence's argument depends upon believing that s_F is more strongly conditioned to entering the food alley, but the argument contains no proof for this belief. Since the rats at the choice point look back and forth, many of the responses of entering the water alley will be directly preceded by a look down the food alley and the response $r_F - s_F$. In this way s_F must become conditioned to turning away from the food alley and entering the water alley. This possibility reveals the need for a proof that s_F is more strongly conditioned to entering the food alley.

(3) Assuming that s_F is conditioned only to entering the food alley, a final question arises. *What reason is there to believe that r_F is more strongly evoked by hunger than r_W?* Clearly this point is essential to the argument, but in discussing it Spence only says: 'see the principle of stimulus dynamism'. This latter principle simply tells us that responses conditioned to hunger will be more vigorous than those conditioned to thirst. But what reason is there for believing that r_F is more strongly conditioned to hunger than r_W? Since reinforcement occurs whenever a hungry rat eats, once a chance r_W becomes conditioned to the hunger stimulus it will never be extinguished, but instead will soon reach maximal habit strength. For this reason Spence's argument needs a proof that r_F is more strongly conditioned to the hunger stimulus than r_W.

III. A MORE COMPLETE r_G EXPLANATION

In this section we shall consider two additional assumptions which will make the r_G explanation of the Kendler experiment more complete.

Let us first assume that *all fractional reactions are incompatible with each other*. Thus if the animal is now making r_F he cannot also be making r_W. Such an assumption seems implicit in Spence's argument when he speaks of 'the interference with r_F to food by

r_W' (Spence, 1951, p. 280), though he never makes this assumption an explicit part of his argument.

Next let us assume that *an r_G is reinforced only if it is simultaneous with an R_G to an S_G*. Now with these two assumptions we can provide proof for the three points discussed in the preceding section.

(1) Since r_G's are reinforced only when they are simultaneous with an R_G to an S_G, the r_G's evoked by the compound drive state will not become conditioned to the alley cues. In fact, the only r_G's which can become conditioned to the alley cues are those that occur in the goal box in the presence of traces from the alley. Since r_F is incompatible with both r_W and R_W, the only r_G's which become conditioned to the water alley are r_W's. It follows then that r_F cannot become conditioned to the water alley.

(2) Since only r_W is conditioned to the water alley and since r_F is incompatible with r_W, when the rat turns from looking down the food alley to looking down the water alley, r_F is replaced by r_W. As a consequence, once the water alley is evoking r_W, s_F cannot immediately precede and accompany responses of entering the water alley. It follows that running down the water alley is more strongly conditioned to s_W than to s_F.

(3) Since R_F is the only R_G which will reduce the hunger drive and produce reinforcement, and since r_F is the only r_G compatible with R_F, it follows that the only r_G which can become conditioned to the hunger drive is r_F. Therefore, it follows that r_F is more strongly conditioned to the hunger drive than r_W.

In Appendix A the r_G argument with these two added assumptions is presented in detail. Though this logical derivation is largely self-explanatory, a few introductory comments may prevent some misunderstandings.

For the sake of simple exposition, I have avoided introducing the complex apparatus necessary to elaborate the generalization gradient of stimulus traces. Instead, I have simply separated those traces which are active at the time of reinforcement from those which are not. A crude distinction like this is all that is required by the argument. Hence the strange distinction between 'supported and unsupported stimulus traces'. It should be clear that the same argument can be formulated with the more precise and more cumbersome notion of a generalization gradient.

Secondly, instead of making a special assumption about the reinforcement of r_G's, I have assumed (see *CR*1) that no R is reinforced unless it is contiguous with an R_G to an S_G. Spence (1947) has already made this point in his analysis of the gradient of primary reinforcement.

Thirdly, I have had to make some precise assumptions about the conditions of training in order to derive precise conclusions. Thus I assume precise values for the life of a stimulus trace (see *D*2 and *D*3), precise values for the number of reinforcements required to condition an R to an S (see *CR*3), and precise values for the running speed of Kendler's rats (see *EC*3 and *EC*4). These values are, of course, quite arbitrary; others would serve as well. What is important is the form of the argument, not these particular assumptions.

Finally, I have saved Spence's distinction between the associative and non-associative functions of the r_G. In Demonstration 1 the associative function is demonstrated, and in Demonstration 2 the non-associative function is demonstrated.

IV. A TEST OF THIS EXPLANATION

Granting that the argument in Appendix A is valid, we have in Appendix B the argument which provides our experimental test of this r_G explanation of the Kendler experiment.

The only difference between the argument in Appendix B and that in Appendix A is that the experimental conditions *EC*8–*EC*10 have been replaced by new assumptions about the experimental conditions. I assume a maze like Kendler's (except that the side which is not black is painted white). Food is placed in *both* goal boxes (*EC*8·5); the rats are hungry but not thirsty throughout *both* the training and test trials (*EC*11); during training the floor in the black section is smooth masonite, while the floor in the white section is covered with hardware-cloth mesh (*EC*9·5); and in the test trials a masonite floor is present in *both* the black and white sections (*EC*10·5). If we now substitute r_{Fb} (eating in the black box) and r_{Fwt} (eating in the white box) for r_F and r_W in the Kendler experiment, we can derive the conclusion that: *when tested with a masonite floor in both the black and white sections of the maze, the rats are now more likely to enter the black alley than the white.* Appendix B demonstrates this conclusion using an argument which is identical in form to the argument in Appendix A. In this sense, then, the following experiment tests our r_G explanation of the Kendler experiment.

(1) *Apparatus, subjects, training.** The maze was like Kendler's with one side, including half of the starting stem and all of the goal box, painted black, while the rest of the maze was painted white. The width of the alleys, the length of the arms, the size of the goal boxes, and the placing of the guillotine doors were the same as in Kendler's maze. There was only one difference. The alley floors of the two halves of the maze were different. Thus in the white side of the maze the floor was covered with hardware cloth ($\frac{1}{2}$ in. mesh) and the black side was smooth masonite. This was true for half of our animals. For the remaining animals the pairing of alley colour and flooring was reversed (see Fig. 3).

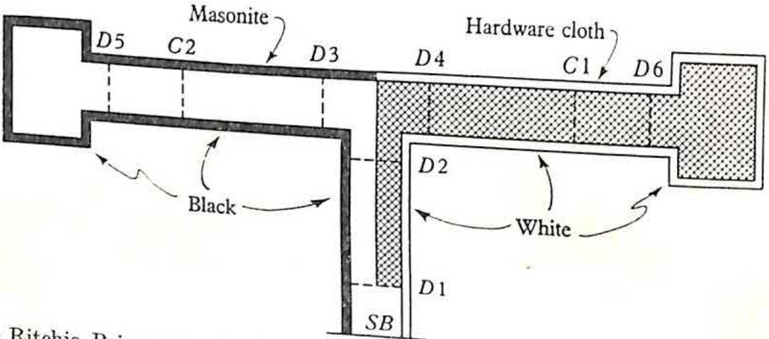


Fig. 3. The Ritchie–Peirce maze during pre-test training. The floor of the maze stem was divided lengthwise, one half of which was masonite, the other half hardware cloth.

Thirty pigmented naïve male rats from the University of California colony were given preliminary handling in which they learned to eat wet Purina mash from bird baths. Throughout the entire experiment they were kept on a 20-hour food deprivation schedule with water available in the home cage at all times.

Each animal was then given four daily trials separated by approximately 5 minutes on 11 successive days. Two of each day's trials were free-choice trials and two were forced following the schedule used in the original Kendler experiment. A black bird-bath with wet Purina mash was placed in the black goal box, and a white one with wet mash in the white goal box. The animals were allowed to eat for 20 seconds following each run. At the end of pre-test training, each animal had made twenty-two rewarded runs down each of the two alleys.

---

* Mr Trevor Peirce, now of Harvard University, trained and tested the animals in the experiments.

(2) *Test trials.* The animals were then matched for alley preferences exhibited in the choice trials. These two groups were then tested in the maze under one of two conditions shown in Fig. 4A and B. The choice in this test was between a white and a black alley, *both* of which now had the same flooring, either masonite or hardware cloth. One group was tested on trials one and four with masonite flooring, and on trials two and three with hardware cloth flooring. The matched group was tested on trials one and four with hardware cloth and on trials two and three with masonite.

If in these tests an animal chose an alley entailed by the argument in Appendix B, he was allowed to eat for 20 seconds. If he chose the opposite alley, he found an empty food container in his goal box. By this device I hoped to provide the degree of secondary reinforcement for 'wrong' choices which Kendler's rats received when they chose the water alley while hungry, and the food alley while thirsty. The test trials were separated from each other by approximately 5 minutes.

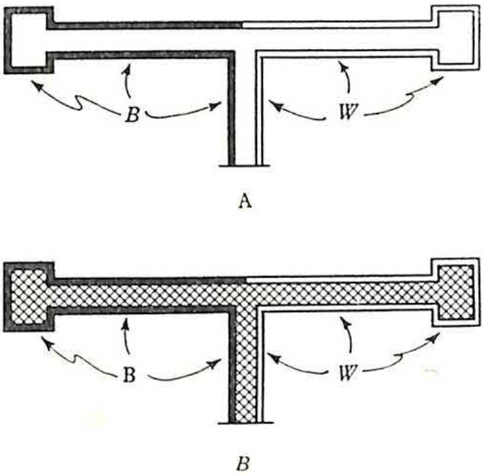


Fig. 4A, B. The Ritchie–Peirce maze during testing. A: Masonite floors in both white and black sides of the maze. B: Hardware cloth floors in both white and black sides of the maze.

(3) *Results of the test trials.* In examining the results of these test trials we shall compare them with those reported by Kendler. Two questions are primary: (1) What was the nature of the alley preference before the test? (2) What was the nature of the alley preference in the two test conditions?

We have already mentioned that in the early pretest training our animals showed a preference for the hardware cloth alley. Though this preference weakened with training, it was still present at the end of the pretest training. However, since the hardware cloth was associated with the white alley for half the animals and with the black for the remainder, and since in the test the hardware cloth appeared either in both alleys or in neither, this preference cannot have biased our test results and so may be ignored. More to the point is an examination of the black–white preference during the pretest training.

On the first test trial half of the animals are required by our hypothesis to choose the black alley and half to choose the white alley. It is appropriate, then, to ask how frequently, in the pretest trials, the black alley was chosen by the first group and how often the white alley was chosen by the second group. In Table 1 we see the number and percentage of animals in the last eight pretest trials that chose the alley which was then made correct on the first test trial. In the last column on the right I have estimated from Kendler's graphical representation of his data the percentage of his animals that chose the alley which was later made correct on the first test trial. The total values reveal that there was during the last 4 days (eight prechoice trials) of pretest training no preference for the 'correct' or for the 'incorrect' alley. The values are very close to those estimated from the Kendler data. In our experiment there was a slight but insignificant preference in both groups for the black alley. This difference, however, was not large enough to require a separate analysis of the test results for these two groups.

Table 1. *The frequency of 'appropriate responses' on free choice trials for four successive days before the first test trial*

	Responses to side made correct on first test trial						
	Black correct test 1		White correct test 1		Total		Kendler*,
Days	N	%	N	%	N	%	%
1	17	57	12	40	29	48	42
2	16	53	15	50	31	52	56
3	19	63	13	43	32	53	60
4	15	50	14	47	29	48	48
Total	67	56	54	45	121	50·5	51·4

* Estimated from graph in Kendler's paper (1946).

Table 2. *Frequency of correct test responses in Kendler* (1946) *and in Ritchie–Peirce*

	Correct responses in four successive test trials											
	Thirsty (Kendler)		Hardware cloth (Ritchie–Peirce)		Hungry (Kendler)		Masonite (Ritchie–Peirce)		Total (Kendler)		Total (Ritchie–Peirce)	
Trials	N	%	N	%	N	%	N	%	N	%	N	%
1	12	92	8	53	5	71	4	26	17	85	12	40
2	7	100	7	46	8	62	3	20	15	75	10	33
3	7	100	8	53	11	85	8	53	18	90	16	53
4	13	100	3	20	5	71	1	6	18	90	4	13
Total	39	98	26	43	29	73	16	26	68	85	42	35

Let us now compare the test results in our experiment with those in Kendler's. Table 2 presents the frequence of correct test responses in the two experiments on four successive trials. It is apparent in this table that while Kendler's animals showed a significant preference for the correct alley on all four test trials, our animals failed to show this preference on any of the four trials. In fact, if we consider the frequency (35%) for all four trials, our animals showed a slight but insignificant aversion for the 'correct' alley.

Finally, in Fig. 5, the results from these two experiments are graphically represented. In this figure we have plotted the frequency, in percentage, of the choices to the side called 'correct' on test trials one and four. These frequencies are plotted for the prechoice trials on the 4 days before the test, and then for the following four test trials. Between Kendler's results and ours there is a close similarity during the pretest trials. But during the test trials this similarity disappears.

Can the failure of our animals to respond 'correctly' be explained by the possibility either that they did not notice the flooring cues during training, or that they did not notice the change in these cues during the test? We have already mentioned that our animals began the pretest trials with a preference for the alleys with hardware cloth flooring. In this sense, they noticed the flooring cues from the outset of training. The fact that they noticed the change in flooring cues is evident in Fig. 6, where we have plotted the median hesitation times at the choice point for the prechoice trials on the four days before the test and on the first two test trials. Furthermore, as Table 3 makes clear, there is no increase in correctness when the hesitation time increases. If anything is true, it is the rats who hesitate the most that make the most errors.

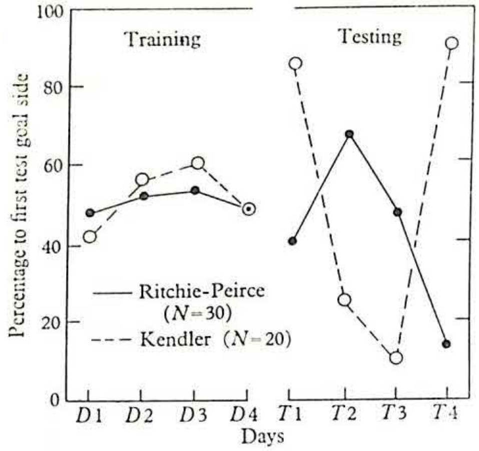


Fig. 5

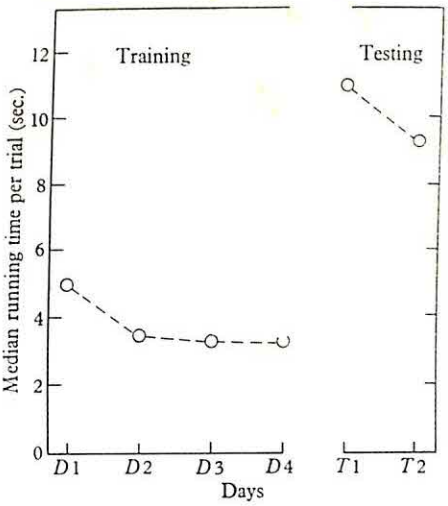


Fig. 6

Fig. 5. Choices in the Kendler and in the Ritchie-Peirce experiments. Percentage of animals during training and testing which chose the maze arm appropriate on test trials $T1$ and $T4$.

Fig. 6. Running time in seconds during training and during the first two test trials.

Table 3. *Frequency of correct responses for groups with different latencies in the test*

Groups	Correct responses in first test N	%
1. S's with latencies less than in training. $N=8$	4	50
2. S's with latencies the same as in training. $N=3$	2	66
3. S's with latencies greater than in training. $N=19$	6	31

V. CONCLUSIONS

The great weakness in the r_G hypothesis, according to Deutsch's (1956) searching analysis, is that there is no proof that the alley-evoked r_G will change when the character of the reinforcement in the goal box is changed. As long as reinforcement follows, Deutsch argues, the ${}_sH_R$ for the alley-conditioned r_G must increase. Our new postulate $P2$, which asserts that all fractional reactions are incompatible, repairs this weakness. Thus when an animal enters a former food box, his conditioned r_F must stop and be replaced by r_W before the animal is reinforced by $S_W \to R_W$. This entails that when water replaces food in the goal box, counter-conditioning will replace r_F in the alley by r_W. Deutsch considers such a change in assumption and argues that even with this change the r_G hypothesis will not explain the Kendler experiment unless we also assume that the rat is stimulated by s_F *only after* he has turned to look down the food alley. But, says Deutsch, 'why should the animal choose to expose itself to one set of stimuli rather than the other? Looking right or looking left will exclude one set of stimuli. However, both sets of stimuli will tend, owing to previous learning, to evoke a response, though one more strongly than the other. If the choice is random, the animal will expose itself to either set of stimuli indifferently. Once one set of stimuli is made to appear in such a way that the other set will concomitantly disappear, the animal's behaviour will be determined by it. Thus, if the animal looks at the weaker set, it does not help to say that the other set would be stronger if the animal looked at it, for it is the set which is impinging which is at that moment stronger' (p. 398).

This is a shrewd objection, but it must be recognized that it is directed not towards the r_G hypothesis but towards the *S-R* analysis of choice behaviour. Without an answer to this objection we cannot understand how an animal can perform properly in any discrimination where the animal is required to turn away from the cue which has less *relative* excitatory strength. If we accept Deutsch's argument, we would never expect the animals to give up their positive habits, and our only evidence of discrimination would be found in the response latencies to the positive and negative cues in the preferred position. Thus the question which Deutsch raises asks us to provide an explanation of *VTE* behaviour, and this is not easy.

What comes to mind first is, of course, Hull's notion of behavioural oscillation, but it is clear that we need further principles to explain why ${}_SO_R$ is greater to the cue which has less excitatory strength. Then there is the further problem of explaining why this tendency to look away from the negative cue persists as the cue's absolute excitatory strength (generalized from the positive cue) increases. The germ of the problem lies in the necessity of basing the oscillation tendency on the absolute and not on the relative excitatory strength of the negative cue, while recognizing that the absolute excitatory strength of the negative cue is constantly increasing. The problem is as important as it is complex, but it is not a problem peculiar to the r_G mediation hypothesis. We can turn then to an examination of the explanatory powers of the r_G.

The evidence from our experiment shows that one or more of the premises in the argument must be false. The most suspect of all our premises is, of course, the assumption that eating in a black box is a different response from eating in a white box. If r_{Fb} and r_{Fwt} are not different and therefore not incompatible, the argument in Appendix B is fallacious and our experiment does not test the r_G hypothesis. When we think of these fractional reactions as *responses* rather than as neural reactions (as in Hebb's analysis), it is hard to imagine how they can differ. This is what makes this premise suspect. But there are important theoretical reasons for retaining this premise if we can.

These reasons become apparent when we apply an r_G analysis to the phenomena of the sensory conditioning and reasoning experiments. In these experiments animals *appear* to learn that stimulus B follows stimulus *A* and that stimulus *D* follows stimulus *C*, though they show no overt response to *B* which is different from the responses made to *D*. The explanatory power of the r_G would be extended considerably if we could believe that to such stimuli as *D* and *B* the animal makes different *reactions*, fractions of which become attached to stimuli which precede them. Then, Brogden's (1939) sensory conditioning experiments could be regarded simply as instances of mediated generalization. Such an explanation is, as Bugelski (1956) points out, 'still a moot question although there is every logical ground to assume some mediating process, perhaps of the type of Hebb's phase-sequence variety. It is not necessary that the mediation occur on an overt, observable, response level. A neural mediation is logically sufficient. Unless we accept some kind of mediating mechanism, we shall be unable to account for the most interesting as well as the most human behavior' (p. 153).

The same kind of theoretical reasons for retaining this premise arise when we try to account for the evidence of reasoning in animals. In Hull's (1935) analysis of the three table problem, the response of stepping on to a platform to drink is taken as equivalent to the response of stepping off this platform to eat. Hull further assumes that drinking on one platform is a different response from drinking on another platform, and that the fractional reactions of these different responses produce different s_G's. All that appears to distinguish different fractional reactions are the stimuli which evoke them at their point of origin. The same idea is implicit in an r_G analysis of Seward's (1949) experiment which demonstrated that rats can latently learn location of empty black and white goal boxes.

For these reasons, it is important to retain this premise until the facts require that it be rejected. Without this premise there is no way that an r_G hypothesis can account for the learning which occurs when the consequences of instrumental responses differ only in their stimulus characteristics.

There are, then, two immediate problems facing the r_G hypothesis. First, we must seek evidence for this 'perceptual learning' premise. Secondly, assuming that the evidence supports the premise, we must consider how best to alter the mediation postulates to account for the evidence of our present experiment. To illustrate what I mean, let me suggest a change in postulate $P4$, which in Appendix A asserts that the K-value of a stimulus is determined by the intensity of s_G. $P4'$: *The greater the similarity between the* r_G *evoked by internal stimuli and the* r_G *evoked by the alley, the greater the* K*-value of* S. This one change in the postulates enables us to account for both our results and Kendler's. In the Kendler experiment, the r_G evoked by the hunger drive is more like the r_G's evoked by the food alley than like the r_G's evoked by the water alley. As a consequence, when hungry the animals chose the food alley and when thirsty, the water alley. In our experiment, on the other hand, the r_G evoked by the hunger drive is no more like the r_G evoked by the black path than like the r_G evoked by the white path. As a consequence our animals show no change in preference when the conditions of testing are changed.

As this paper has proceeded, we have altered Spence's r_G hypothesis more and more until now the idea is very close to Seward's (1956) recent conceptions. We have changed it by introducing the following new premises: (1) that all fractional reactions are incompatible: (2) that fractional reactions are reinforced only when contiguous with an R_G to an S_G: (3) that reactions to different stimulus patterns have different fractional reactions; and (4) that the incentive motivation is determined by the similarity between the drive-evoked r_G's and the externally evoked r_G's. Such a neo-Spencian r_G hypothesis will account for the facts, while the classic Spencian hypothesis will not.

Let me conclude by comparing this neo-Spencian hypothesis with the classic Tolman interpretation of purposive behaviour. For Tolman the problem has two parts: (1) the animal in traversing the maze alley learns what leads to what; and (2) when put at the choice point, the animal chooses that alley which leads to what he now wants.

This neo-Spencian r_G hypothesis shows one sense in which the animal may learn what leads to what. That is to say, the r_G–s_G at the alley entrance has characteristics which are uniquely determined by what the animal in the past has found in the goal box. Secondly, the new postulate $P4'$ shows one sense in which the animal chooses the alley which leads to what he wants. That is to say, the alley which evokes an r_G which is more like his internally evoked r_G is the alley with the greater incentive motivation.

What difference, if any, is there, then, between the classic Tolman hypothesis and our neo-Spencian hypothesis? Pragmatically considered, the two cannot be distinguished. But in another sense there is a difference between the two theories. Spence and Hull began with a premise that the elements of behaviour are S's and R's and that the only kinds of changes which occur in the structure of behaviour are changes in the associative connexions between S's and R's. Given this methodological premise, their task was to devise a mechanism to account for purposive behaviour within the restrictions of this premise.

Tolman, on the other hand, has from the very first been sceptical of this premise, largely because of the vagueness in the original conceptions of S and R reinforcement. Thus, instead of constructing from these elements a mechanism to account for purposive behaviour, Tolman has directed his attention to evaluating the failures of the S-R reinforcement hypothesis, searching among the ruins for a key to the proper solution to the problem. As yet, no such key has been found.

Meanwhile, the S-R theorists, largely in reply to the experimental criticisms of Tolman and other cognitive theorists, have modified their conceptions of S and R and reinforcement to an extent where we now discover in the neo-Spencian hypothesis an associationistic analysis of purposive behaviour in which reinforcement is largely secondary, the crucial R's are fractional reactions possibly wholly neural, and the crucial S's are the sensory consequences of such reactions! All that is left of the original programme is the belief that the required principles of association can be isolated and studied in simple conditioning experiments where S's and R's and reinforcements are palpable and definite. Given these experimentally grounded principles and a set of postulates about fractional reactions, we can, as we have seen, devise a testable explanation of purposive behaviour. This, I think, is the key for which cognitive theorists have been searching.

If this is true, this marks the end of the controversy between the S-R and cognitive theories. To the S-R theorists goes the credit for having seen how purposive behaviour could be explained in an associationistic analysis. To the cognitive theorists goes the credit for exposing the weaknesses in the earlier S-R reinforcement conceptions, and for suggesting, primarily through their experiments, what kinds of changes were required in these fundamental conceptions.

Let me close by emphasizing that when I say that this controversy is ended, I do not mean that all the theoretical differences which have divided these two groups of psychologists are now resolved. Nor do I mean that we now have a final analysis of purposive behaviour. All I mean is that on this one issue of the difference between an 'expectancy' interpretation and an 'r_G' interpretation of purposive behaviour, there can no longer be any basic difference. What must now be done is to devise a more precise mediational analysis of such behaviour, and here I foresee future controversies even more exciting than the old one.

VI. APPENDIX A

Definitions

*D*1. A stimulus trace is supported during interval I (following the termination of S_x) if and only if the stimuli during I are physically like S_x.

*D*2. A supported stimulus trace remains *active* for at least 2·0 sec.

*D*3. An unsupported stimulus trace decays in 0·5 sec. and becomes *inactive*.

$D4$. The only R_G for a hungry rat is R_F, eating food, and the only R_G for a thirsty rat is R_W, drinking water.
$D5$. The only R_G's for hungry and thirsty rats are R_F and R_W.
$D6$. A fractional reaction is *terminal* if and only if it is part of an R_G to an S_G.
$D7$. An R is conditioned to an S if and only if S acts to evoke R whenever S or its trace is active.
$D8$. The greater the K-value of an S, the greater is the likelihood of S acting to evoke whatever reactions are conditioned to it.

Postulates

$P1$. Each R_x to S_x has one and only one fractional component, r_x.
$P2$. No two fractional reactions are compatible.
$P3$. Each fractional reaction, r_x, produces its own stimulus pattern, s_x.
$P4$. The greater the intensity of an s_G simultaneous with an S, the greater is the K-value of the S.

Conditioning rules

$CR1$. A reaction is reinforced to an S if and only if S or its trace is active when R occurs, and R is contiguous with an R_G to an S_G.
$CR2$. When two or more stimuli act to evoke a common reaction, the resulting reaction is more probable and more vigorous than when they act to evoke incompatible reactions.
$CR3$ After an S-R pair is reinforced six times, the R is conditioned to the S.

General theorems

$GT1$. When an R_G to S_G occurs, the only contiguous fractional reaction is r_G, the fractional component of R_G. ($P1$, $P2$.)
$GT2$. When an R_G to S_G occurs, the only contiguous r_G is the terminal r_G. ($D6$, $GT1$.)
$GT3$. A fractional r_G is reinforced if and only if it is a terminal r_G. ($CR1$, $GT2$.)
$GT4$. The reaction r_F but not r_W is conditioned to the hunger stimulus. ($D4$, $GT3$.)
$GT5$. When two or more stimuli act to evoke a common r_G, the resulting reaction will be more probable and more vigorous than when they act to evoke different r_G's. ($P2$, $CR2$).

Experimental conditions

$EC1$. All SB traces are supported while running from SB to the black goal box.
$EC2$. No SB traces are supported while running from SB to the unpainted goal box.
$EC3$. All rats reach a goal box within 2·0 sec. after leaving the choice point.
$EC4$. No rat reaches a goal box within 0·5 sec. after leaving the choice point.
$EC5$. Each alley response ends as R_G, the goal response, begins.
$EC6$. All alley responses are compatible with R_G.
$EC7$. Before the test the rats are given more than twenty-four training trials.
$EC8$. Only the black goal box contains food, and only the unpainted goal box contains water.
$EC9$. During the training the rats are both hungry and thirsty.
$EC10$. During the test the rats are hungry but not thirsty.

The argument in theorem form

$T1$. All SB traces in the black goal box are active. ($D2$, $EC1$, $EC3$.)
$T2$. No SB traces in the unpainted goal box are active. ($D3$, $EC2$, $EC4$.)
$T3$. Terminal r_W's occur only in the unpainted goal box. ($D5$, $D6$, $EC8$.)
$T4$. No terminal r_W is accompanied by an active SB trace. ($T2$, $T3$.)
$T5$. Each terminal r_F is accompanied by an active SB trace. ($T1$, $T3_b$.) [Note: Wherever a theorem is given a b subscript, this refers to an exactly comparable theorem about the S-R relations in the opposite side of the maze.]
$T6$. SB is not active whenever r_W is reinforced. ($GT3$, $T4$.)
$T7$. No SB-r_W pairs are reinforced. ($CR1$, $T6$.)
$T8$. Each reinforced r_F is accompanied by an active SB trace. ($GT3$, $T5$.)
$T9$. Some SB-r_F pairs are reinforced. ($CR1$, T8.)
$T10$. After twelve trials r_F but not r_W is conditioned to SB ($CR3$, $T7$, $T9$.)
$T11$. After twelve trials each RB is preceded and accompanied by s_F but not by s_W. ($D7$, $P3$, $T10$.)

T 12. Each *RB* is reinforced during training. (*D* 5, *EC* 5, *EC* 6, *EC* 9.)
T 13. After twenty-four trials *RB* but not *RU* is conditioned to s_F. (*CR* 1, *CR* 3, $T\,11_{ab}$, *T* 12.)
T 14. In the test *SB* and hunger act to evoke r_F but not r_W, and *SU* acts to evoke r_W but not r_F. (*D* 7, *GT* 4, *EC* 7, *EC* 10, $T\,10_{ab}$.)
T 15. In the test the vigour of r_G is greater when a rat is stimulated by *SB* than when stimulated by *SU*. (*GT* 5, *T* 14.)

Demonstration 1

To prove: The rats in the test are more likely to make *RB* than *RU*.
1. In the test more stimuli act to evoke r_F than r_W. (*T* 14.)
2. *RB* but not *RU* is conditioned to s_F. (*EC* 7, T 13.)
3. The stimuli in the maze act to evoke *RB* but not *RU*. (*D* 7, *CR* 2.)
Therefore: etc.

Demonstration 2

To prove: The rats in the test are more likely to make *RB* than *RU*.
1. In the test the intensity of s_G simultaneous with *SB* is greater than the intensity of s_G simultaneous with *SU*. (*P* 3, *T* 15.)
2. The *K*-value of *SB* is greater than the *K*-value of *SU*. (*P* 4.)
3. The stimuli in the maze act to evoke *RB* but not *RU*. (*D* 8.)
Therefore: etc.

VII. APPENDIX B

Definitions. (The same as in Appendix A.)
Postulates. (The same.)
Conditioning rules. (The same.)
General theorems. (The same.)

Experimental conditions

EC 1–*EC* 7. (The same.)
EC 8·5. Food is present in both goal boxes during training.
EC 9·5. During training *SM* (masonite flooring) is present in the black section of the maze, and *SH* (hardware cloth flooring) in the white section of the maze.
EC 10·5. During the test *SM* is present in all sections of the maze.
EC 11. During training and test the animals are hungry but not thirsty.

The argument in theorem form

T 1–*T* 13. (The same, given the appropriate substitutions of *SWt* for *SU*, and $r_F b$ for r_F, and $r_F wt$ for r_W.)
T 10·5. After twelve trials $r_F b$ but not $r_F wt$ is conditioned to *SM*. (*CR* 1, *CR* 3, *GT* 1, *EC* 9, *T* 3.)
T 13·5. After twenty-four trials *RB* but not *Rwt* is conditioned to *SM*. (*CR* 1, *CR* 3, *EC* 9, *T* 12.)
T 14·5. In the test *SB* and *SM* act to evoke $r_F b$ but not $r_F wt$, and *SWt* acts to evoke $r_F wt$ but not $r_F b$. (*D* 7, *EC* 10, *T* 10, *T* 10·5.)
T 15·5. In the test the vigour of r_G is greater when the rat is stimulated by *SB* than when stimulated by *SWt*. (*GT* 5, *T* 16.)

Demonstration 1

To prove: The rats in the test are more likely to make *RB* than *RWt*.
1. In the test more stimuli act to evoke $r_F b$ than $r_F Wt$. (*T* 14·5.)
2. *RB* but not *RWt* is conditioned to $s_F b$ and to *SM*. (*EC* 7, *EC* 10·5, *T* 13, *T* 13·5.)
3. The stimuli in the maze act to evoke *RB* but not *RWt*. (*D* 7, *CR* 2.)

Demonstration 2

To prove: The rats in the test are more likely to make *RB* than *RWt*.
1. In the test the intensity of s_G simultaneous with *SB* is greater than the intensity of s_g simultaneous with *SWt*. (*P* 3, *T* 15·5.)
2. The *K*-value of *SB* is greater than the *K*-value of *SWt*. (*P* 4.)
3. The stimuli in the maze act to evoke *RB* but not *RWt*. (*D* 8.)
Therefore: etc.

REFERENCES

BROGDEN, W. J. (1939). Sensory pre-conditioning. *J. Exp. Psychol.* **25**, 323–32.

BUGELSKI, B. R. (1956). *The Psychology of Learning.* New York: Holt.

DEUTSCH, J. A. (1956). The inadequacy of the Hullian derivations of reasoning and latent learning. *Psychol. Rev.* **63**, 389–99.

HULL, C. L. (1935). The mechanism of the assembly of behavior segments in novel combinations suitable for problem solutions. *Psychol. Rev.* **42**, 219–45.

KENDLER, H. H. (1946). The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. Exp. Psychol.* **36**, 212–20.

SEWARD, J. P. (1949). An experimental analysis of latent learning. *J. Exp. Psychol.* **39**, 177–86.

SEWARD, J. P. (1956). Drive, incentive and reinforcement. *Psychol. Rev.* **63**, 195–203.

SPENCE, K. W. (1951). Theoretical interpretations of learning. In *Comparative Psychology*, ed. by C. P. Stone. New York: Prentice-Hall.

SPENCE, K. W. (1947). The role of secondary reinforcement in delayed reward learning. *Psychol. Rev.* **54**, 1–8.

(*Manuscript received* 12 *August* 1957)

QUANTITATIVE JUDGEMENT IN SOCIAL PERCEPTION*

By H. TAJFEL

University of Oxford Delegacy for Social Administration

Stimuli which form a series may differ from each other in their physical magnitudes and in their value to the subject; they may also be classified according to various attributes. A theoretical schema is presented which is based on some assumptions about the effects of the interaction of these three variables upon absolute judgements pertaining to series of stimuli. These assumptions allow for predictions of shifts of judgement in a number of series in which these variables are found in various combinations. Many of these combinations are representative of conditions under which judgements in social situations are made. An attempt is also made to apply some predictions following from the schema to judgements involving not only physical but also abstract continua.

I. Physical environment in social perception

The main concern of the study of social perception is with the ways in which people perceive, judge, and interpret the social objects and events which together amount to what is called their 'social environment'. There is also, however, a great number of investigations in this field which could be grouped together as being concerned with the effects of social environment on the perception, judgement, and interpretation of the physical environment. Anthropologists have often been concerned with differences between cultures in the nature and amount of sensory information required to infer the existence of a physical phenomenon; for example, see Hallowell's (1951) account, in which the belief in the physical presence of a supernatural being is based predominantly on the interpretation of a sequence of auditory cues. Similarly, visual symbols used for purposes of representation or communication from the artist to his public have varied considerably from one historical period to another; for example, see Wittkower (1955). It is very possible that the psychological explanation of these differences (which could by no means be self-sufficient) is to be sought in the fact that social consensus provides an important and repeatedly validated part of the information which an individual receives about his environment. There is quite a gap between these phenomena, which involve some of the most basic aspects of a culture, and the modest perceptual departures from the objectively measurable physical properties of stimuli which psychologists have been investigating in their laboratories. However, some similarities can also be found.

These laboratory studies sought to determine the effects of social standards, pressures, and valuations on the perception or judgement of some selected physical properties of the environment. Many of them were concerned with shifts in the perception or judgement of, for example, size, weight, or colour, which followed in their direction the pull exercised by the various social factors. Verbal labels have been shown to exercise an influence of that nature: Baker & Mackintosh (1955) recently reported that labelling of meaningless shapes of identical shade with names of colours influenced the way in which the colour of these shapes was perceived at a later stage when they were presented again without the verbal labels. Linguistic labels have been shown to be effective in a variety of other situations. Their 'social' aspect is not in the fact that they represent something which is different in principle from many other cases where past experience intervenes in the judgement, perception, reproduction, or interpretation of the physical environment. It is rather in the fact that in these instances past experience is crystallized in, and transmitted through, the social medium of language.

---

* Based on papers read at the Annual Conference of the British Psychological Society in St Andrews, March 1957, and at the XVth International Congress of Psychology in Brussels, August 1957.

Similar shifts have been observed in the investigations concerned with the effects of group standards on judgements of magnitudes of stimuli. Tresselt (1948) found discrepancies between groups of weight-lifters and watch-makers in the initial judgements of heaviness of a series of weights. Here again, as Tresselt points out, there is nothing specifically 'social' in the processes responsible for the shifts, apart from the fact that the differences in the past experience of the contrasted groups of subjects were determined by membership of different professional groups. In the same way, a statement that a car is 'big' or 'small' may have very different meanings, according to the side of the Atlantic on which it is made. Assessment (which can be quantified) of the nature of the past experience with the stimuli provides a basis from which these shifts in judgements can be predicted, independently of whether this experience has been gained in a social context.

The same is true of shifts in judgements of magnitude induced by group pressures which are active at the time when the judgements are made. Experiments by Asch (1951) and by Bovard (1951) can serve as examples here. Judgements of physical magnitudes made by members of groups have been shown, in these experiments, to shift, sometimes quite drastically, in the direction of the group consensus. As in the previous examples, these phenomena can be referred for explanation to considerations of a general nature about the learning and motivational processes active in the situation. Thus, group consensus is generally effective in inducing shifts because it is one of the sources, consistently reliable in normal circumstances, of information about the environment. The effectiveness of the pressures may vary from individual to individual, as in Asch's (1951) experiments, or from group to group, as in the results reported by Bovard (1951). These individual and group differences reflect differences in the degree of conformity which may depend in turn on the nature and extent of the emotional involvement of the individual with the group of which he is a member. It is quite possible these experiments represent a modest and very simplified version of the processes involved in the cultural and historical differences mentioned earlier.

All these examples of the influences exercised by the social environment on the judgement and perception of the physical environment share two characteristics: independently of their social character, they can be discussed in terms of general judgemental or perceptual processes; and they can all be conceived as 'unidirectional' shifts in judgement or perception. As was pointed out earlier, the direction of the discrepancies between the physical and the perceived or judged properties of the stimuli is determined in all these cases by the direction in which the socially derived standards, pressures, or labels work. There exists, however, another group of findings in which the shifts induced by factors of a social nature cannot be called 'unidirectional' in the same sense, and have not as yet been satisfactorily explained in a way which relates them to a wider theoretical background.

II. JUDGEMENTAL SHIFTS WITHIN SERIES

Shifts in judgements of physical magnitudes reported in this second group of findings can be characterized as *intra*serial, in contrast to those discussed so far; they consist of changes in the judged relationships between the magnitudes of stimuli *within* a series. One example of such intraserial shifts is to be found in the accentuation of perceived or judged differences between the stimuli of a series where value changes concurrently with the physical magnitude which the subjects are requested to judge, as in the case of series of coins; another, in the fact that prejudiced individuals tend to accentuate the differences in some physical characteristics between people who are members of the group against which they are prejudiced, and those who are not (Secord, Bevan & Katz, 1956).

The accentuation of perceived differences between the stimuli of a series probably accounts for a number of phenomena in the field of perceptual over-estimation (cf. Tajfel, 1957, for a more detailed discussion). It is important to note that in most experiments on over-estimation, and in those concerned with the effects of prejudice on judgements of physical characteristics, data were collected by presenting the stimuli individually. This has been done either by asking the subjects to assign quantitative verbal labels to the stimuli presented one by one, or by asking them to match a variable standard to each stimulus in turn. There are good reasons for maintaining that these methods are adequate for investigating a number of genuine perceptual phenomena. As Bruner (1957) has recently argued, it is very possible that most perceptual activities 'depend upon the construction of an adequate system of categories against which stimulus inputs can be matched' (p. 127). He continues as follows:

> There is probably one condition where perceptual acts are relatively free from such influences, and that is in the task of discriminating simultaneously presented stimuli as alike or different....Ask the person to deal with one stimulus at a time, to array it in terms of some magnitude scale, and immediately one is back in the familiar territory of inferential categorizing. Prentice, in his able defense of formalism in the study of perception, seems to assume that there is a special status attached to perceptual research that limits the set of the observer to simple binary decisions of 'like' and 'different' or 'present' and 'absent', and to research that also provides the subject with optimal stimulus conditions, and Graham has recently expressed the credo that no perceptual laws will be proper or pure laws unless we reduce perceptual experimentation to the kinds of operations used in the method of constant stimuli....(But) the point must be made that many of the most interesting phenomena in sensory perception are precisely those that have been uncovered by departing from the rigid purism of the method of constants (pp. 127–8).

The main assumption discussed in the paper referred to above (Tajfel, 1957) was that in a series of stimuli 'where value changes concurrently with the dimension subjected to investigation, the differences between the stimuli of the series will be perceived as larger than the objectively equivalent differences between the stimuli of a neutral series, where no such association exists between value and magnitude' (p. 193). One of the supporting arguments for this assumption was found in the existing evidence that, in a series of quantitative judgements pertaining to a series of stimuli, judgements of any one individual stimulus are affected by all other stimuli of the series. The initially rather vague concept of the frame of reference has recently acquired a precise meaning in the field of judgement. We are now able to predict how the response scale will shift as a function of changes in the range and frequency of presentation of the various stimuli of the series, of their magnitudes, of the segment of the series at which anchoring is introduced, and of the extent to which the stimuli may or may not be perceived as belonging to the series which is being judged. Judgements of size obtained in the experiments on over-estimation, in which the subjects had to deal continuously with series of stimuli, cannot be understood through an exclusive consideration of the relationship between a response and an individual stimulus to which it relates. They must be seen in the light of the background relationships between the stimulus being judged and all other stimuli of the series to which it belongs.

There is a good deal of evidence (summarized in Tajfel, 1957) to support the view that this serial nature of the situation in the over-estimation experiments, coupled with the intervention of the variable of value, resulted in the accentuation of perceived differences between the stimuli. This accentuation is best understood as an improvement in the acuity of discrimination between the stimuli of the series. The fact that what is here called an improvement in discrimination resulted also in a discrepancy between the actual and the perceived magnitudes of the stimuli (i.e. over-estimation) need not be considered as a contradiction in terms. The term 'discrimination', as applied to judgements of series of stimuli, is not free of ambiguities. There are at least four different criteria for assuming that an improvement in discrimination has taken place: increased number of recognizable steps in the series; decrease in the variability of judgements pertaining to the individual stimuli; closer correspondence of judgements with the physical values; and increased perceived distance, as measured by the distances between the mean assignments of values to the stimuli, between the elements of the series. This last criterion is not necessarily the worst. As Stevens (1957) writes, the '...philosophy of indirect measurement asserts that all we can know about magnitude is what confusion tells us. Variability becomes the measure of things, and the mean is meaningless' (p. 154). But the mean can be meaningful in two ways at least: it may tell us how close is the correspondence between the quantitative judgement and the actual magnitude of the stimulus; and it may also tell us what happens to the perceived, as compared with the actual, relationships between the elements of a series. An improvement in discrimination may consist simultaneously in a closer correspondence between the actual and perceived values of the stimuli, *and* in increased distances between the values assigned to the neighbouring stimuli; but it may also be conceived in terms of increased distances between these values, leading to a departure from the correspondence with the actual values, and reflecting the fact that the stimuli of the series are perceived as more different from each other than they really are.

This is what seems to happen, under conditions of judgement outlined above, in series which display concomitant changes in the value of the stimuli to the subject and in the physical magnitudes of the stimuli (see Tajfel, 1957, for an argument attempting to account for this phenomenon). Series of coins, which have somehow taken an undue prominence in all this field of research, provide one example. Another instance may be found, as mentioned earlier, in the perception by prejudiced individuals of some physical features of objects of their prejudice. Here again, the intraserial relationships, in other words the perceived or judged differences between the elements of a series, come to the fore.

The experiment by Secord *et al.* (1956) can serve as a starting-point for further discussion. They reported that a group of prejudiced subjects accentuated the differences in skin colour (and in some other physical features) between negroes and whites more than a group of non-prejudiced subjects. The situation can be described in the following manner: the physical characteristic, the skin colour, is a dimension which varies continuously from light to dark. On this continuous variation a classification is superimposed: the distinction between whites and negroes. Skin colour is one of the determinants of this classification; in other words, there is a fairly consistent relationship between the discontinuous classification and the continuous physical dimension. So far, the situation is identical for the prejudiced and the non-prejudiced subjects. The difference between the two groups of subjects can be defined only in terms of the emotional, or if one prefers the term, value relevance of the classification to the subjects. It can be said that the negroes and whites are more different from each other for the prejudiced than for the non-prejudiced individuals. The continuous dimension of skin colour is broken by the discontinuous classification which, in the case of prejudiced subjects, is more sharply accentuated. As in the case of studies on coins, the object of judgement is perceived for what it is, in this instance negro or white, before judgements of skin colour are made. Prediction can be made to the effect that judgements on any dimensions which are in some way relevant to a value classification by providing cues to it (and thus acquire value in themselves) will show a shift in opposite directions for stimuli falling into the distinct classes. The situation is not, in principle, different from that encountered in a series of coins. Instead of dealing with individual stimuli which differ concurrently in a physical attribute and in the attribute of value, we have here groups of stimuli which differ in precisely the same manner.

III. Statement of predictions

A value differential, a classification, and a physical dimension enter into a specific relationship in the example just discussed. This particular combination of attributes is, of course, only one of many possible combinations. Before proceeding to a more detailed discussion of these, the predictions concerning the judged relationships between the elements of the various series which result from the combination of attributes should be stated more formally.

(1) When a variation in value is correlated in a series of stimuli with a variation in a physical dimension, the judged differences in this physical dimension between the elements of the series will be larger than in a series which is identical with regard to the physical magnitudes of the stimuli, and in which the stimuli do not possess the attribute of value.

(2) When differences in value exist in a series of stimuli, but are in no way correlated with the variation in a physical dimension, these differences in value will have no effect on the judged relationships between the physical magnitudes of the stimuli of the series.

(3) When a classification in terms of an attribute other than the physical dimension which is being judged is superimposed on a series of stimuli in such a way that one part of the physical series tends to fall consistently into one class, and the other into the other class, judgements of physical magnitudes of the stimuli falling into the distinct classes will show a shift in the directions determined by the class membership of the stimuli, when compared with judgements of a series identical with respect to this physical dimension, on which such a classification is not superimposed.

(4) When a classification in terms of an attribute other than the physical dimension which is being judged is superimposed on a series of stimuli, and the changes in the

physical magnitudes of the stimuli bear no consistent relationship to the assignment of the stimuli to the distinct classes, this classification will have no effect on the judged relationships in the physical dimension between the stimuli of the series.

(5) When a classification in terms of an attribute other than the physical dimension which is being judged is superimposed on a series of stimuli in such a way that one part of the physical series tends to fall consistently into one class, and the other into the other class, and this classification is of inherent value or of emotional relevance to the subject, judgements of physical magnitudes of the stimuli falling into the distinct classes will show a shift in directions determined by the class membership of the stimuli when compared with judgements of a series identical with respect to this physical dimension, on which such a classification is not superimposed; this shift will be more pronounced than the shift referred to in (3) above.

(6) When a classification in terms of an attribute other than the physical dimension which is being judged is superimposed on a series of stimuli, this classification being of inherent value or of emotional relevance to the subject, and the changes in the physical dimension bear no consistent relationship to the assignment of the stimuli to the distinct classes, this classification will have no effect on the judged relationships in the physical dimension between the stimuli falling into the distinct classes. (One exception to this statement will be discussed later.)

IV. APPLICATION OF PREDICTIONS TO THE VARIOUS SERIES

Table 1 contains a list of the various possible series with a brief description of their characteristics. They can now be considered in more detail.

(*a*) *The* pv *series*

This series has already been discussed in a previous section of the paper. Series of coins are an example. Not only are the judged differences between the stimuli larger in *pv* than in *p* series, but there are also some indications that, when the differences in value between the stimuli can be quantitatively assessed, a larger difference in value between the stimuli leads to a greater accentuation of size differences between them. A conversion of the data from the experiments by Bruner & Goodman (1947), Carter & Schooler (1949), and Bruner & Rodrigues (1953) shows that the percentage increase from the actual to perceived differences between the sizes of coins is very much smaller for the difference between a 'penny' and a 'nickel' than between a 'nickel' and a 'quarter'. (See Table 2; a 'penny' is one cent, a 'nickel' 5 cents, and a 'quarter' 25 cents.) As can be seen from Table 2, the percentage increases from the actual to the perceived distances tend to become smaller for all control stimuli in all three experiments as the larger end of the series is approached. They tend to become considerably larger in the case of coins. It is very possible that the over-all differences in percentages between the three experiments are due to differences in the experimental procedure; but the general trend of relative increase in the perceived differences between the coins with the increase in the differences in value between them appears uniformly in all the three studies. The possibility that this is due to factors other than value, such as differences in design between the coins, must not be overlooked. This seems, however, unlikely in view of the finding, recently reported by Vroom (1957), that judgements of size of 'defaced' coins (i.e. with their design

27.8.70 J233

practically removed through use) did not differ from judgements of coins with their design intact. Still other factors may play a role; but there is here at least a suggestion which warrants further investigation.

Table 1

Series	Characteristics of the series
1. p	Ordered change in a physical dimension
2. pv	Ordered change in a physical dimension correlated with a change in value
3. pc_1	Ordered change in a physical dimension consistently related to a classification in terms of another attribute
4. pc_2	Ordered change in a physical dimension; a classification in terms of another attribute superimposed on the series bears no consistent relationship to the change in the physical dimension
5. $(pv)c_1$	Ordered change in a physical dimension correlated with value; a classification in terms of another attribute consistently related to the change in the physical dimension is superimposed on the series
6. $(pv)c_2$	Ordered change in a physical dimension correlated with value; a classification in terms of another attribute superimposed on the series bears no consistent relationship to the changes in physical dimension
7. $p(c_1v)$	Ordered change in a physical dimension is consistently related to a classification in terms of another attribute; this classification is of inherent value or of emotional relevance to the subject
8. $p(c_2v)$	Ordered change in a physical dimension; a classification in terms of another attribute superimposed on the series bears no consistent relationship to the change in the physical dimension; this classification is of inherent value or of emotional relevance to the subject
9. $(pv)\,(c_1v)$	Ordered change in a physical dimension correlated with value; a classification in terms of another attribute superimposed on the series is related to change in the physical dimension: this classification is of inherent value or of emotional relevance to the subject
10. $(pv)\,(c_2v)$	Ordered change in a physical dimension correlated with value; a classification in terms of another attribute superimposed on the series bears no consistent relationship to change in the physical dimension; this classification is of inherent value or of emotional relevance to the subject
11. $(pv_1)\,(c_2v)$	Ordered change in a physical dimension correlated with value in one class of stimuli, not correlated with value in another class; a classification in terms of another attribute superimposed on the series bears no consistent relationship to change in the physical dimension; this classification is of inherent value or of emotional relevance to the subject

Table 2

		Percentage increase from actual to perceived differences between stimuli	
Experiment	Stimuli	Penny–nickel	Nickel–quarter
Bruner & Goodman	Coins	86·3	131·0
	Disks	36·3	20·7
Carter & Schooler	Coins	36·4	53·4
	Aluminium disks	22·5	19·3
	Cardboard disks	24·0	12·1
Bruner & Rodrigues	Coins (on table)	55·5	83·4
	Metal disks	52·7	40·0
	Paper disks	50·0	16·9

(*b*) *The* pc_1 *series*

In the preceding series value and a physical dimension—two of the three variables considered here—were seen in a specific relationship. In the present series, the relationship is between a physical dimension and a classification (see Table 1, series 3). The differences between the stimuli in the physical dimension along which judgements are made are relevant to the classification since they provide cues for assigning the stimuli to one class or another; but they are not the primary basis of the classification. The classification itself may be based on another physical attribute, on an abstract attribute, or on a combination of several attributes. When the stimulus is identified as belonging to one of the classes before the judgement of its magnitude in the physical dimension is made, this class identification serves in turn as a cue to the judgement of the physical magnitude, and causes the shift of judgement in the direction consistent with the general relationship between the physical magnitudes and the classification. The non-prejudiced subjects in the experiment by Secord *et al.* (1956) are confronted with such a pc_1 situation. The judgement of the skin colour of a man labelled as 'white' should, according to the present argument, differ consistently from the judgement of the skin colour of the same man labelled as 'negro'. This is one example of a situation arising frequently in the judgement of physical characteristics of stimuli in the social environment. Classifications of people into ethnic, or sometimes other, groups often imply consistent differences between the groups in terms of some physical features. Industrial or agricultural products originating from different sources may be known to differ consistently in their size, colour, texture, weight, etc. The shifts of judgements could be demonstrated by comparing judgements pertaining to such series with those of other series, identical with the pc_1 series with regard to the physical dimension, and not broken by a discontinuous classification (p series). A comparison of a pc_1 with a pv series should yield similar results.

One further hypothesis can be formulated concerning what might be called the 'zone of uncertainty' between two classes of stimuli in a pc_1 series. Class identification of a stimulus in this series, when it precedes an absolute judgement of the physical magnitude, has so far been considered as capable of exercising a shift in this judgement, since it acts as a cue helping to characterize the object of judgement as possessing the physical attribute to a greater or lesser degree. However, this effect cannot be expected to be of the same extent along the entire range of stimuli. The darker the skin of a negro, or the lighter the skin of a white man, the more redundant is the information provided by the pre-existing classification, and the more will the judgement be based on the actual physical characteristics of the stimulus. This information will, however, increase in importance in the more ambiguous cases towards the middle of the range, such as those of relatively light-skinned negroes and dark-skinned whites. It is here that the largest shifts, consistent with the classification, may be expected.

(c) *The* pc_2 *series*

In such a series there is no orderly relationship between the physical dimension judged and the classification; the physical magnitude of a stimulus does not help in its assignment to one of the various classes, and the identification of a stimulus as belonging to one of the classes does not provide any cue as to its physical magnitude. In this situation, the classification is not expected to have any effects on judgements of physical magnitude. Size of books and their classification into the subjects with which they deal may be taken as one example. There is no consistent relationship between the two, and one would not expect a book to be judged larger or smaller because it is known to be concerned with physics, biology, or psychology. Judgements of physical magnitudes of stimuli belonging to a pc_2 series should thus not differ from judgements relating to a simple p series, and consequently they should differ from pv and pc_1 series in the same way as a p series differs from them; i.e. differences between the stimuli in a pv series should be more accentuated than the corresponding differences in a pc_2 series; and the shifts in opposite directions in the judgements of stimuli belonging to different classes, predicted for a pc_1 series, should not occur in a pc_2 series.

It will be obvious that pc_1 and pc_2 series, as characterized above, are no more than two extremes of a continuum, running from a perfect correlation between an orderly progression in the physical dimension and a classification, as in the case of a pc_1 series, to a complete lack of correlation in the case of a pc_2 series. If it can be assumed that the shifts in the judgements of physical magnitudes are determined by a long-standing relationship between the physical dimension and the classification, then a further hypothesis can be formulated to the effect that the extent of these shifts will be a function of the degree of consistency of this relationship, experienced in the past. Thus, the correlation between skin colour and the classification into whites and negroes is presumably very high, and consequently noticeable shifts in judgements of skin colour may be expected; on the other hand, it may well be true that the Scandinavians tend to be taller, on the average, than the Italians, but as the overlap between the two classes is bound to be very large, shifts in judgements of height may either be non-existent or hardly noticeable.

This hypothesis has some relevance to the general problem of stereotyping, as it formulates a testable relationship between the strength of a stereotype and the nature of past experience which supports it. The fact that stereotypes are essentially consequences of sharpened or accentuated classifications has not been sufficiently exploited in the laboratory. Judgements of almost any aspects of objects which are stereotyped are not made in a vacuum, they are always implicitly comparative. The statement that Italians are 'lively' means that they are more lively than some other people. The problem is to find out to what extent an Italian is judged as being 'lively' (i.e. more lively than somebody else) not because he is lively, but because he is known to the judge to be an Italian. An investigation of this problem must proceed along two lines: first, it is possible to create experimental conditions in which the amount of correlation between a discontinuous classification and a continuous dimension, along which judgements are made, can be varied; secondly, shifts in judgement due to the effects of the various degrees of consistency of this relationship can be investigated by presenting the subjects with identical stimuli variously characterized at different times as belonging or not belonging to a particular class.

(*d*) *The* (pv)c$_1$, p(c$_1$v), *and* (pv)(c$_1$v) *series*

In these series there is an interaction between a change in a physical dimension, a classification, and value. The three series will be treated together, as it is very doubtful whether they can be empirically distinguished, though they may differ in their origin. In the $(pv)c_1$ series there is a correlation between value and a physical dimension, as in the previously discussed pv series. The physical dimension is in turn consistently related to a classification, as was the case in the pc_1 series. The example of industrial or agricultural products originating from two sources, which was used above for pc_1 series, applies here when a change in the quality of the products is related to a change in the physical dimension, such as size, colour, or fineness of texture. When products from one of the sources are consistently found at the favoured end of the physical range, a $(pv)c_1$ series will ensue. The effects on judgements of the physical dimension could be assessed by comparing judgements given when the origin of the items present is unknown with judgements given when each of the items is labelled according to its origin.

The $p(c_1v)$ series presents a slightly different case. The classification is here one of value, and the changes in the physical dimension, not themselves related to value, are related to the classification in a manner described previously for the pc_1 series. As in both the $(pv)c_1$ and the $p(c_1v)$ series two of the terms are correlated with middle third term, the first two terms may also be correlated. Or, it can be said that in the case of the $(pv)c_1$ series the classification which is related to a change in the physical dimension has come to be associated with value, as the change in the physical dimension had such an association in the past. In the $p(c_1v)$ series, the change in the physical dimension would come to be associated with value by a similar process in which the initial association of the classification with value would play the intermediary role.

These two series would thus converge empirically towards a (pv) (c_1v) series, in which value is associated both with the classification and with the change in the physical dimension. Skin colour and other physiognomic traits of negroes and whites, as judged by the prejudiced group of subjects in the experiment by Secord *et al.* (1956), are examples of such series; it is probable that the degree of darkness of skin, or of flatness of nose would not have had a value connotation for these subjects if possession of these attributes in various degrees had not been consistently associated with the classification into negroes and whites which is for them of inherent value. It has been argued above that such an association tends to shift in opposite directions judgements of the 'relevant' physical characteristics of the stimuli falling into the distinct classes. In a corresponding pc_1 classification these shifts were attributed to the fact that the classification, which has built up specific expectations, provides information concerning the physical magnitudes, especially in the middle part of the range. Here, the classification, which is the same, should exercise similar effects; but these effects should be further accentuated as the classification, or the identification of an individual as negro or white, is decidedly more important for the prejudiced than for the non-prejudiced subjects. This special attention paid by prejudiced individuals to physical cues, which has for its purpose the efficient identification of members of the group against which they are prejudiced, has been demonstrated in other contexts by Allport & Kramer (1946), and by Lindzey & Rogolsky (1950). In the context of judgement, Stevens & Galanter (1956) have recently argued that 'what happens when O rates an apparent magnitude on an N-point scale is the result of the interplay of three important "forces" plus an unknown number of lesser ones. The three important forces that interact are: intent, discrimination, and expectation' (p. 7). By 'intent' Stevens and Galanter mean that 'a properly instructed O tries to make the intervals on his category scale equal in width and thereby produce a linear scale' (p. 7). Discrimination is the subjects' ability 'to tell one magnitude from another' (p. 7). Expectation plays a role since the subject 'inevitably has expectations regarding how the experimenter will distribute the stimuli, and consequently regarding the relative frequency with which he should use the various categories' (p. 7). The differences between judgements of skin colour given by prejudiced subjects, and judgements of magnitude given in a 'neutral' psychophysical experiment are in the different ways in which the factors of intent and of expectation influence judgement in these two situations. In the former case, expectation is not determined by the subjects' guesses about the experimenter's distribution of the stimuli falling into the various categories; it is based on the long-standing past experience of a consistent relationship between the classification and the various categories in terms of which judgements are made. And it is not the subject's intent to produce a linear scale: intent can be replaced by the 'intention' of producing a scale whose close correspondence to the classification represents a perceptual 'vested interest'.

Shifts in opposite directions of judgements of the stimuli falling into the distinct classes should thus be larger in a (pv) (c_1v) series than in a corresponding pc_1 series. Some evidence of this is provided by the fact, mentioned above, that the prejudiced subjects perceived negroes as more negroid in physiognomic traits than did the non-prejudiced subjects. In this case, the photographs were first identified by the subjects as belonging to the 'negro' or 'white' category, and the judgements of 'negroidness' were subsequent to this identification. A study by Pettigrew, Allport & Barnett (1958) provides some very interesting indications about the converse phenomenon: the assignment of stimuli to various classes based on ambiguous cues from the physical dimensions which are correlated with a value classification. The experiments were conducted in South Africa; subjects from the five South African ethnic groups (Afrikaans-speaking Europeans, English-speaking Europeans, Coloureds, Indians, and Africans) were asked to identify the 'race' of faces after 'a brief stereoscopic presentation of pairs of racial photographs'. These pairs were of two kinds: either the same face presented to both eyes, or ethnically mixed pairs in all possible combinations. The differences between the Afrikaners and all the other groups are of interest here. (1) For the 'one-race pairs' this group gave the largest percentages of 'European' and 'African' judgements, and the smallest of 'Coloured' and 'Indian' ones; the Europeans and the Africans representing, of course, the extremes of the range. (2) An analysis of the distribution of incorrect identifications for the 'one-race pairs' revealed that the Afrikaners gave fewer incorrect identifications as 'Coloured' and 'Indian' than all other groups, and twice as many 'African' judgements as all other groups combined. (3) For the 'two-race pairs', when responding to the combination with an identification of ethnic group which did not correspond to either of the two presented photographs, the Afrikaners gave the smallest percentage of 'Coloured' responses, and the largest percentage of 'African' responses. Pettigrew *et al.* conclude that 'throughout the results, the Afrikaners deviate most consistently in their judgements from the other four groups. They tend to respond in an either-or, bifurcated manner.' If it is assumed that a racial classification in terms of 'white' and 'non-white' has a more pronounced connotation for the Afrikaner group than for the other groups (including the English-speaking European group), these results are entirely consistent with the above discussion of the (pv) (c_1v) series.

Predictions concerning the differences between the (pv) (c_1v) series and the p and pc_2 series need not be detailed here, as they would be similar to those outlined above in the comparison between these two series and the pc_1 series.

No prediction can be made at this stage concerning the direct comparison of a (pv) (c_1v) series with a pv series, as this raises the problem of the interaction, within one series, between the variable v and a c_1 classification. There is the possibility that a classification of this nature leads to a decrease in the efficiency of discrimination between the stimuli *within* each of the classes; on the other hand, v should lead to an increase in the judged differences between the stimuli along the entire range. The effects of the resulting complex interaction may either cancel each other, or produce differences between pv and $(pv)\,(c_1v)$ similar to those expected when p is compared with $(pv)\,(c_1v)$.

(*e*) *The* (pv)c_2 *series*

One of the main implications of the preceding discussion was that a c_2 classification, which bears no relationship to the changes in the physical dimension, has generally no effects on the judged relationships between the physical magnitudes of objects which fall into the distinct classes. Thus, a $(pv)c_2$ series should not differ in this respect from a pv series; for example, judged differences in size between coins are not likely to be affected by a classification of the coins according to the date of their minting, whether or not the subjects are informed of the date for each coin before they give their judgements of size.

(*f*) *The* p(c_2v) *series*

The above implication concerning the c_2 classification applies also to this series. Some supporting evidence can be found in experiments on perceptual over-estimation which yielded negative results. One example is provided by the experiment of Klein, Schlesinger & Meister (1951) in which judgements of size of disks bearing the sign of swastika were compared with similar judgements of disks bearing 'neutral' signs. No over-estimation of the swastika disks was found, and the authors' conclusion was that this went some way towards invalidating other results in this field. On the assumptions outlined here, the results of this experiment *had* to be negative. It is difficult to conceive of a consistent relationship between the size of a swastika and the degree of its emotional relevance. Therefore, a series of disks of various sizes bearing the emblem is, in the terminology adopted here, a p series and not a pv series in which results similar to those found in the experiments on coins could be expected. In the experiment by Klein *et al.* (1951) the subjects were presented with the swastika and other disks in random order; this resulted in a combination of two p series which together amount to a $p(c_2v)$ series. In this series, there was no association between size and value, there was a value classification into emotionally relevant and emotionally neutral stimuli, and the stimuli belonging to the two classes were, as far as their size was concerned, randomly distributed along the series.

(*g*) *The* (pv) (c_2v) *series*

This is a combination of the two preceding series. It is introduced here only for the sake of completeness, as concrete examples of it are difficult to find; it could, however, be produced without much difficulty in the laboratory. Predictions concerning this kind of series are difficult to make in the context of the present argument, and possibly of little interest, as they are neither crucial to the discussion nor readily applicable in 'real-life' situations. Difficulty of prediction is due to the fact that if the value classification does not bear any relationship to the progression in the physical magnitude, the random distribution of stimuli with and without the attribute of value along the physical dimension would necessarily destroy the pv aspect of the series, i.e. the correlation between value and the physical dimension.

(*h*) *The* (pv_1) (c_2v) *series*

This series is introduced as it provides an experimental counterpart to the previously discussed $p(c_2v)$ series. It will be remembered that the prediction concerning the latter series was that it should not differ from a simple p series, and that this prediction found some support in the negative results of studies on perceptual over-estimation. The difference between the present series and the $p(c_2v)$ series lies in the fact that in one of the two classes of stimuli presented for judgement here there is a correlation between value and magnitude; there is no such correlation in the other class of stimuli. A series of coins, considered on its own, is a pv series. When it is combined with other series in experiments in which coins and other disks are presented for judgement in alternation, the result is the (pv_1) (c_2v) series. Its characteristics are as follows: there is a classification in terms of value into coins and other disks; in terms of size, there is a random distribution of stimuli from both classes along the series; and the association of size with value applies to coins only, not to the other stimuli. As a result, the effects of a pv situation, which is in the case of coins a product of long-term past experience, may be expected to apply to coins only, and not to the other stimuli. The nature of these effects has already been discussed.

V. ABSTRACT CONTINUA

It is possible that the schema outlined so far may be applicable not only to the effects of various kinds of classifications on judgements of physical magnitudes, but also to their effects on abstract judgements which imply the existence of a continuous attribute. This kind of judgement can hardly be considered 'perceptual'; but there are no *a priori* grounds to assume that principles which are found helpful in the prediction of judgements made under certain conditions, and concerned with the physical aspects of stimuli, should not be capable of application to judgements of abstract attributes made under similar conditions. Many objects in our environment, especially in the social environment, are often rated along some quantitative or quantifiable abstract continuum, such as beauty, pleasantness, intelligence, etc. Such 'dimensions' are often, by their very nature, inherently possessed of, or correlated with, value, and they may be, in this sense, compared to the previously discussed pv series. On this abstract equivalent of a pv series discontinuous classifications in terms of other attributes may be superimposed in a manner similar to those intervening in a pv series. Such classifications may or may not be of inherent value to the subject. If a series of paintings, half of which are Dutch and half Italian, are presented for rating in terms of beauty to someone who has no previous bias in favour of either, the fact that the subject knows which paintings are Dutch and which Italian should not affect the results. This is a near enough equivalent of the previously discussed c_2 classification, when it is considered that beauty is now the continuum along which judgement is made, just as the physical dimension was before. If, on the other hand, the classification in terms of another attribute is itself of value, the judgemental results should be similar to those discussed above under the headings of c_1 and c_1v classifications. In an experiment on stereotypes, Razran (1950) asked his subjects to rate photographs of faces in terms of their pleasantness and other similar attributes. The same photographs were presented later, together with a number of new ones, and ethnic labels were attached to each one of them. The new judgements, as compared with the original ones, tended to show displacements which might be expected on the basis of a value classification originating in prejudice. This is, of course, an example of a familiar phenomenon which occurs in an infinite variety of social situations. It is introduced here because it relates theoretically to the preceding argument, and because predictions of effects on the basis of this argument could be made in each case in which the attribute determining the value classification could be identified.

REFERENCES

ALLPORT, G. W. & KRAMER, B. M. (1946). Some roots of prejudice. *J. Psychol.* **22**, 9–39.

ASCH, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (ed.): *Groups, Leadership and Men*. Pittsburgh: The Carnegie Press.

BAKER, K. E. & MACKINTOSH, I. (1955). The influence of past associations upon attributive colour judgments. *J. Exp. Psychol.* **49**, 281–86.

BOVARD, E. W. (1951). Group structure and perception. *J. Abnorm. Soc. Psychol.* **46**, 398–405.

BRUNER, J. S. (1957). On perceptual readiness. *Psychol. Rev.* **64**, 123–51.

BRUNER, J. S. & GOODMAN, C. C. (1947). Value and need as organizing factors in perception. *J. Abnorm. Soc. Psychol.* **42**, 33–44.

BRUNER, J. S. & RODRIGUES, J. S. (1953). Some determinants of apparent size. *J. Abnorm. Soc. Psychol.* **48**, 17–24.

CARTER, L. F. & SCHOOLER, K. (1949). Value, need, and other factors in perception. *Psychol. Rev.* **56**, 200–7.

HALLOWELL, A. I. (1951). Cultural factors in the structuralization of perception. In Rohrer, J. H. & Sherif, M. (eds.): *Social Psychology at the Crossroads*. New York: Harper.

KLEIN, G. S., SCHLESINGER, H. J. & MEISTER, D. E. (1951). The effect of values on perception: an experimental critique. *Psychol. Rev.* **58**, 96–112.

LINDZEY, G. & ROGOLSKY, S. (1950). Prejudice and identification of minority group membership. *J. Abnorm. Soc. Psychol.* **45**, 279–86.

PETTIGREW, T. F., ALLPORT, G. W. & BARNETT, E. O. (1958). Binocular resolution and perception of race in South Africa. *Brit. J. Psychol.* **49**, 265–78.

RAZRAN, G. (1950). Ethnic dislikes and stereotypes: a laboratory study. *J. Abnorm. Soc. Psychol.* **45**, 7–27.

SECORD, P. F., BEVAN, W. & KATZ, B. (1956). The Negro stereotype and perceptual accentuation. *J. Abnorm. Soc. Psychol.*, **53**, 78–83.

STEVENS, S. S. (1957). On the psychophysical law. *Psychol. Rev.* **64**, 153–81.

STEVENS, S. S. & GALANTER, E. H. (1956). Ratio scales and category scales for a dozen perceptual continua. *U.S. Navy, Project Nr.* 142–201, *Report PNR*–186.

TAJFEL, H. (1957). Value and the perceptual judgment of magnitude. *Psychol. Rev.* **64**, 192–204.

TRESSELT, M. E. (1948). The effect of the experience of contrasted groups upon the formation of a new scale of judgment. *J. Soc. Psychol.* **27**, 209–16.

VROOM, V. H. (1957). Design and estimated size of coins. *Canad. J. Psychol.* **11**, 89–92.

WITTKOWER, R. (1955). Interpretation of visual symbols in the arts. In *Studies in Communication*. London: Secker and Warburg.

(*Manuscript received 5 September* 1957)

ATTENTION TO VISUAL DISPLAYS DURING A VIGILANCE TASK

II. MAINTAINING THE LEVEL OF VIGILANCE

By C. H. BAKER

Applied Psychology Research Unit, Medical Research Council, Cambridge

Analysis of data from an earlier paper by the writer (Baker, 1958) shows that perceptual performance does not necessarily deteriorate during a typical vigilance task. It was hypothesized that a major determinant of such deterioration is the regularity of signal appearance. An experiment is described here which supports this hypothesis. It is further demonstrated that motor activity markedly increases during a vigilance task, the least active subjects tending to detect more signals.

I. INTRODUCTION

It has generally been recognized that in tasks of the vigilance type, perceptual performance deteriorates in time. The classical clock-test data of Mackworth (1944, 1948, 1950) have been repeatedly confirmed (Adams, 1956; Andrews, Ross & Whittenburg, 1956; Bakan, 1955; Pepler, 1953). Another aspect of the vigilance situation has been noted by Wallis (personal communication, 1955). Wallis found that subjects tended to become more 'restless' as the vigilance task progressed. In so far as the level of vigilance generally shows deterioration in time, and motor activity ('restlessness') increases, it seemed not improbable that the level of vigilance might be predicted from a measure of motor activity. In designing the following experiments, therefore, an attempt was made to quantify motor activity. Baker (1958) recently described a series of experiments on attention to visual displays over a period of 64 min. His report dealt with the manner in which visual attention was distributed over displays. The present paper is concerned with the vigilance aspects of the experiments.

II. METHOD

The method has been described by Baker (1958). Briefly, subjects sat facing a circular display of light 12 in. in diameter on which a pip (a small spot of light) might appear at any location. Actually, pips appeared twice at each of sixteen locations in random order. The temporal interval between consecutive appearances ranged from 36 to 196 sec., the average inter-pip period being 2 min.

In the first experiment the display was blank, except for the occasional appearance of a pip. In the second and subsequent experiments the display was radar-like in that a revolving radial line (6 r.p.m.) 'painted on' the pips, which lasted for 1 sec. In the third experiment a small light above the display comes on at irregular intervals for 2, 3, or 4 min., and the subjects were directed to search for pips in the area behind the outer half of the radial line only, when the light was on. In the fourth experiment only the outer half of the line was displayed, while in the fifth a revolving 'box-sweep' was employed rather than the single line.

When the subject was sitting upright in the fixed chair, the viewing distance to the display was 18 in. Upon seeing a pip, the subjects' task was to press a response button, and response time was measured. Subjects' movements were recorded on telephone counters activated by microswitches placed under each corner of the chair seat, the seat being balanced one eight of an inch above the chair frame on a central pivot.

III. Results (experiments 1–5)

In four of the five experiments there was *no* progressive deterioration in the percentage of pips missed during four consecutive 16 min. intervals (Fig. 1). (The decrement between 16 and 48 min. for the Half-sweep experiment is not significant.) The exception was the Green-light experiment, the decrement between 16 and 64 min. being significant at the 0·02 level of confidence. Later, in the discussion, attention is drawn to the markedly different experimental conditions that may have accounted for this exception. The original curve obtained by Mackworth (1944), which shows a significant decrement between first and second half-hour periods in a 2-hr, task, has been plotted on Fig. 1 for comparison.

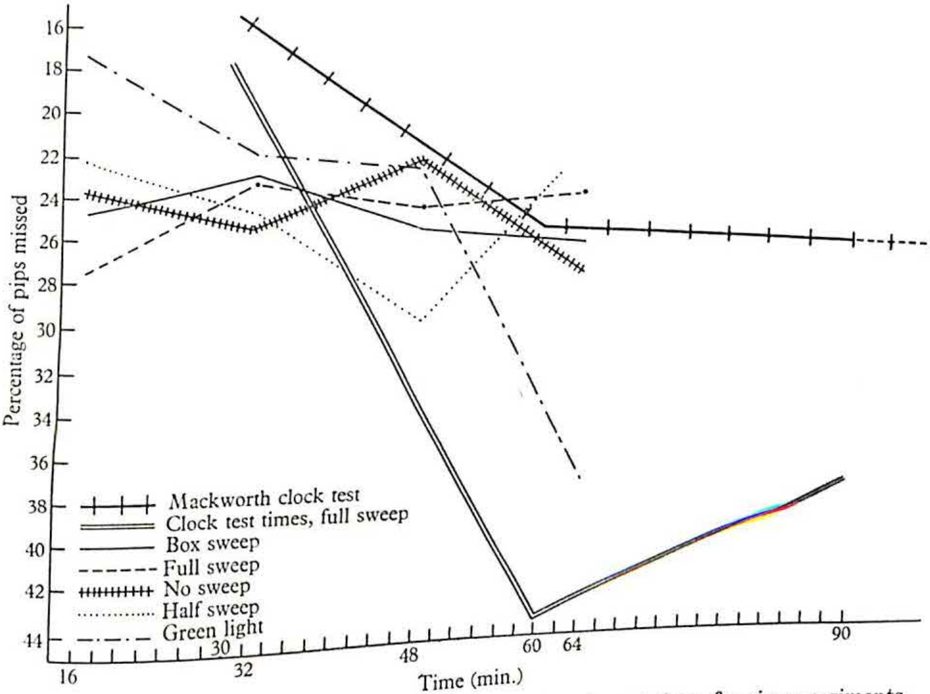


Fig. 1. Showing the percentage of pips missed, as a function of time, for six experiments. Mackworth's data (1944) have also been plotted for comparison purposes.

It appeared that the major difference between the above experiments and Mackworth's was the interval between signal appearance. Whereas in these experiments fifteen pips appeared each half-hour, on an average of one every 2 min., Mackworth employed twelve signals each half hour. Mackworth's signals were clustered into the first 20 min. of the half-hour, followed by a blank period of 10 min. (The difference of three signals per half-hour is considered of negligible importance. An analysis of the performance of the 'worst' and 'best' eight subjects in each of the five experiments showed that in the four 'no-decrement' experiments the 'worst' subjects were representative of the whole group in each case, i.e. also showed no significant decrement in time. And for these subjects the apparent rate of signal appearance was actually less than for Mackworth's subjects.) To test our hypothesis that the regularity of signal appearance was a major factor in determining whether vigilance would deteriorate in time, the present experiment was performed to duplicate the time intervals between signals which Mackworth employed.

Experiment 6—*Full-sweep* (*Clock-test intervals*)

Twelve of the sixteen pip locations used in the previous experiments were employed—nos. 1, 2, 3, 6, 7, 8, 9, 11, 12, 13, 14, 16, a different location in each azimuth quadrant being ignored. Intervals between pip appearance were, as for Mackworth, $\frac{3}{4}$, $\frac{3}{4}$, $1\frac{1}{2}$, 2, 2, 1, 5, 1, 1, 2, and 3 min., followed by a blank spell of 10 min. to complete the half-hour. On successive half-hours this pattern was repeated, the experiment running for $1\frac{1}{2}$ hr. The conventional Full-sweep display (experiment 2) was employed. Subjects were thirty naval ratings, each of whom had been a subject in one of the first five experiments at least one week previously. As for Mackworth, all subjects performed the task in the same order, i.e. all had the 10 min. blank period in the last 10 min. of each of the three half-hour intervals.

IV. Results (experiment 6)

The results are shown superimposed on Fig. 1. A very marked deterioration of performance in time was shown, the difference between performance during the first and second half-hour periods being significant at the 0·001 level of confidence. The increment in performance during the third half-hour is not significant. In addition, 19% of the pips escaped detection and attention was distributed in the familiar pattern revealed by the previous experiments. In summary it can be said that the regularity of pip appearance is a major factor in determining whether or not vigilance will deteriorate during a period of an hour.

V. Results (motor activity)

Because some subjects were far more active than others it was decided to show the increase in activity with time in terms of the percentage of total activity for each 16 min. interval, an arbitrary scale, rather than in terms of counter readings. The data from experiment 6 (Clock-test intervals) have been presented for both 64 and 80 min. periods. These data are shown in Table 1. From this table it is evident that motor activity progressively increased with time, being about $2\frac{1}{2}$ times greater in the last quarter of the task than in the first quarter. Furthermore, all experiments were remarkably consistent in this respect. Table 2 shows the degree of correlation (Spearman rank-difference) between motor activity scores and number of pips which were not detected. Though it is apparent from Table 2 that the relation between motor activity and performance on the task was significant in only three of the six experiments, it should be noted that the non-significant figures are all positive. Presumably more or all of them would have been

Table 1. *Showing the percentage of motor activity* (*arbitrary scale*) *during each* 16 *min. interval, for six experiments*

Experiments	Intervals (min.) 0–16	16–32	32–48	48–64	64–80	Total percentage
No-sweep	14·9	20·9	31·9	33·1	—	100
Full-sweep	11·1	22·2	32·3	34·4	—	100
Green-light	12·9	25·5	27·5	34·1	—	100
Half-sweep	12·1	25·7	29·0	33·2	—	100
Box-sweep	10·2	19·6	27·7	42·5	—	100
Clock-times (64 min.)	15·4	26·1	28·0	30·5	—	100
Clock-times (80 min.)	11·5	19·6	21·0	22·9	24·88	100

significant had N been larger in each case. In summary it can be said that there is little possibility of predicting the level of vigilance from motor activity during the task. On the other hand, motor activity during a vigilance task increases markedly in time, the less active persons tending to achieve superior performance.

Table 2. *Showing correlation (rho) between motor activity scores and number of targets missed*

Experiment	N	rho	Significance level
No-sweep	16	0·47	0·05
Full-sweep	16	0·39	N.S.
Green-light	16	0·30	N.S.
Half-sweep	16	0·54	0·05
Box-sweep	16	0·16	N.S.
Clock-times	30	0·57	0·01

VI. DISCUSSION

A great variety of factors have been suggested as the cause of the oft-obtained decrement in performance during vigilance tasks. Thus Mackworth (1948, 1950) favours a theory of internal inhibition. Bakan (1957) proposes a two-factor theory; a desire to get to sleep brought on by the monotony of the environment (the sleep factor), conflicting with the desire to please the experimenter, resulting in self-stimulation to remain awake. The more one is self-stimulating, the more targets are missed. Broadbent (1953) has found that performance deteriorates in a paced situation, but not in one which is unpaced, and ascribes the reason to increasing lapses in attention. Bowen (1956) credits a lack of visual noise on the display as a major determinant of decrement.

Environmental factors such as high temperature (Pepler, 1953) and auditory noise (Broadbent, 1953) have been found detrimental, as has sleep deprivation (Wilkinson, personal communication, 1955). Factors which have been found to minimize the effect include visual noise (Bowen, 1956), the presence of the experimenter in the subject's room (Fraser, 1953), rest periods, knowledge of results, external stimuli, and benzedrine (Mackworth, 1950), and intense signals (Adams, 1956; Mackworth, 1950). Each of these investigators employed different temporal patterns of target appearance.

It is not generally made clear that the vigilance problem is really dichotomous; on one hand is the problem of average level of vigilance throughout the task, while on the other is the problem of the deterioration of vigilance in time. A recent paper by Deese (1955) and the work of Bowen (1956) firmly establish the fact that the proportion of targets detected (or the number of times a target must blink before being detected) is a function of the number of targets presented. Thus Deese found that at ten targets per hour less than 50 % were detected, while for forty per hour the percentage was nearly 90. (We must mention in passing Wallis's note (Anonymous, 1956) that in both cases the number of targets missed was the same—about four per hour.)

Deese examined two alternative hypotheses, the reinforcement versus the expectancy hypothesis, in seeking an explanation of his results. Briefly, the reinforcement hypothesis states that, 'the probability of detection for any given signal is determined exclusively by the temporal location of the signal immediately preceding the signal in question'. The expectancy hypothesis states that, 'the probability of detection—is determined by a large and rather indeterminant number of signals preceding the signal in question—the observer—continuously performing a kind of averaging of previous input in order to extrapolate the results to future behaviour of the search field'.

The reinforcement hypothesis predicts that the shorter the time interval between two signals, the greater the probability of the second being detected. An analysis of Deese's data led him to reject this hypothesis. An analysis of our data from experiments 1 to 5 resulted in a similar rejection, i.e. there is no relation between length of time-interval preceding the appearance of a pip and the probability of that pip being detected, the intervals varying, as stated above, from 36 to 196 sec.

With reference to the expectancy hypothesis Deese has deduced that, 'expectancy should be low immediately after a signal, should increase as the mean intersignal interval is approached, and finally should become quite high as the intersignal interval grows beyond the mean'. But surely this is a truncated deduction—we must add that as the interval grows still longer expectancy *must* again fall to a low level. Granted this additional deduction and another by Deese that, 'if the task is invariant with time, vigilance will change little after the average level of expectancy has been reached', then the experiments above offer confirmatory evidence for the expectancy hypothesis.

The major difference between the first five experiments and the sixth (Clock-times) was, intentionally, the regularity of the intervals between pip appearance. In experiments 1–5 pips appeared on an average of once every 2 min., never closer than 36 sec., never further apart than 196 sec. In the sixth experiment the average rate was similar—once every 2½ min., but the range was relatively enormous—from 45 to 645 sec. It is convenient here grossly to quantify this irregularity between intervals in terms of the ratio of shortest to longest. For experiments 1–5 the ratio was 1:5·4, while for experiment 6 it was 1:14·3. In four of the first five experiments there was no decrement in time. In the sixth the decrement was, as for Mackworth, most marked. Thus we are led to conclude that these experiments have supported the two deductions stated above. In experiments 1–5 (excluding experiment 3) the task was 'invariant with time'. By 'invariant' we mean that pips appeared while expectancy was at or near its peak as determined by the degree of relative regularity of the preceding pips. However, when the marked irregularity of the Clock-times (14·3) was introduced the subjects, 'continuously performing a kind of averaging of previous input', must inevitably have built expectancies which were not reinforced, resulting in expectancy dropping again to a low level. As the number of such non-reinforced expectancies accumulated, performance deteriorated.

It remains to examine this notion of 'continuous averaging'. There is evidence tc suggest that humans are surprisingly adept at this task. Hake & Hyman (1953) have demonstrated that subjects can indeed perceive statistical structuring in a sequence of events and reproduce this structuring in their own series of responses. According to them subjects don't perceive statistical structuring as such, i.e. they don't perceive the probability rules governing the generation of the series of events. Instead, they perceive the sequences of events which precede each prediction and provide information concerning the future behaviour of the series. 'Probability learning' on the basis of the preceding sequence of events has also been demonstrated by Jarvik (1951) and Goodnow (1955). (As further evidence of human sophistication in this area Ross (1955) has shown that mathematically naive subjects can construct a mathematically random binary series of symbols, contrary to a commonly held belief.) We, of course, are primarily interested in 'continuous averaging' of *temporal* series. In discussing time perception Woodrow (1951) has stated that the most important characteristics of the stimulus series underlying the ability to perceive temporal grouping are, 'the relative intensities of the members of the series, their duration,—and their temporal spacing'. And Woodrow presents data showing that it is more difficult to estimate long time intervals than short ones.

Other data are available to reinforce our general position. Bowen (1956) studied the effect of signal frequency on the general level of vigilance. To his surprise he found no decrement in performance over a period of an hour and was led to conclude, 'the presence of the video-noise on the display as the major determinant of our subjects maintained performance', in contrast to Mackworth's (1950) and Bakan's (1955) demonstrations of decrement in which they employed 'monotonously invaried' situations. Bowen used a radar-like display without a sweep-line to investigate signal frequencies of one, ten, and twenty signals per hour, the signals blinking until detected. An examination of Bowen's experimental records shows that at twenty signals per hour, the intersignal interval was *absolutely regular*, appearing at 2, 5, 8, 11, 14, etc. minutes—a ratio of 1:1. He employed three sequences for the condition of ten per hour, the smallest/largest ratio being 1:3·5, 1:2·8, and 1:2·8. Finally, in the condition of one signal per hour, 'it would therefore seem to be immaterial where, in an hour's vigilance, a lone signal occurs'.

Concerning Bowen's 'noise' explanation of results, it should be pointed out that in the present experiments no noise was employed and in four of the first five no decrement was obtained. Indeed, it is difficult to see how the presence of noise would affect more than the general level of vigilance. Noise was purposely avoided in the present experiments for two reasons: (1) it is known that noise level affects pip detectability (Williams, Bartlett & King, 1948) and so a statement that noise was employed is rather meaningless unless the noise can be quantified—a difficult task in simulated situations, and (2) the presence of noise leads to errors of commission (false reports of pip appearance)—there were none in these experiments.

With respect to experiment 3 (Green-light) we must point out that this experiment differed from the remainder of the first five in one important respect. Here the subject was given a task to which he was just warming up when the nature of the task was suddenly changed. Such pacing must inevitably result in successive attempts at 'continuous averaging' which must be abandoned before expectancies can be built. Indeed, with the nature of the search task changing every 2, 3, or 4 min., the number of intervals available for 'averaging' would range from none to two. In support, Broadbent (1953) has demonstrated that when 'a 5-choice reaction task was set to subjects under the same conditions as a vigilance task', it was found that, 'in the paced condition the output of correct responses fell off after 10 minutes work, while in the unpaced condition there was no decrement in output till nearly an hour had gone by'.

Finally, it is necessary to consider motor activity. Bowen actually observed his subjects without their knowledge and noted, every 5 min., 'what the subject had been doing'. He found, 'very little change occurred during the test period', though individuals varied considerably, being 'restless throughout', or 'consistently motionless'. He concluded, 'that subjects who are restless tend to perform poorly'. Our data indicate a marked increase in motor activity in time, which, presumably, cannot be detected by casual observation. However, we agree with Bowen that 'restless' subjects tend to be poor performers. Why this is so is not clear. Bakan (1957) believes that such activity is a form of self-stimulation generated to reduce the soporific effect of the monotonous situation.

With incidental reference to the commonly held view that vigilance decrement results from 'monotony', it should be pointed out that the experiments reported here were excessively 'monotonous'; a revolving line rotating at a constant speed in a circular field of unchanging brightness, while auditorily the only sound heard was a continuous and unchanging noise of 58 decibels.

VII. Conclusions

During a period of an hour there is no deterioration in the level of vigilance if the events searched for are fairly regularly spaced in time. Temporal irregularity between such events and pacing are major factors in producing deterioration. In addition, motor activity during a vigilance task increases markedly in time, least active subjects tending to be superior performers on the task.

This study was undertaken while the author was on a tour of duty with the Medical Research Council, from the Defence Research Medical Laboratories, Defence Research Board of Canada, in Toronto. It is a pleasure to acknowledge the constructive criticisms of this draft by Dr Paul Bakan, Dr R. B. Bromiley, and Dr N. H. Mackworth.

References

Adams, J. A. (1956). Vigilance in the detection of low-intensity signals. *J. exp. Psychol.* **52**, 204–8.

Andrews, T. G., Ross, S. & Whittenburg, J. A. *Perceptual and Motor Skills*, **6**, 109–16.

Anon. (1956). Vigilance—the nature of alertness and the problem of its maintenance during long spells of work. *Nature, Lond.*, **178**, 1375–7.

Bakan, P. (1955). Discrimination decrement as a function of time in a prolonged vigil. *J. exp. Psychol.* **50**, 387–90.

Bakan, P. (1957). Vigilance. *The Advancement of Science*, **13**, 389–93.

Baker, C. H. (1958). Attention to visual displays during a vigilance task. I. Biasing attention. *Brit. J. Psychol*, **49**, 279–88.

Bowen, H. M. (1956). On the appreciation of serial displays. Ph.D. Thesis, University of Cambridge.

Broadbent, D. E. (1950). The twenty dials tests under quiet conditions. *A.P.U. Report No.* 130/50.

Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.* **44**, 295–303.

Broadbent, D. E. (1953). Classical conditioning and human watch-keeping. *Psychol. Rev.* **60**, 331–9.

Deese, J. (1955). Some problems in the theory of vigilance. *Psychol. Rev.* **62**, 359–68.

Fraser, D. C. (1953). The relation of an environmental variable to performance in a prolonged visual task. *Quart. J. exp. Psychol.* **5**, 31–2.

Goodnow, J. J. (1955). Response sequences in a pair of two-choice probability situations. *Amer. J. Psychol.* **68**, 624–30.

Hake, H. W. & Hyman, R. (1953). Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.* **45**, 64–74.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in a serial anticipation of alternative symbols. *J. exp. Psychol.* **41**, 291–7.

Mackworth, N. H. (1944). Notes on the clock test—new approach to the study of prolonged visual perception to find the optimum length of watch for radar operators. *Med. Res. Council Report No.* 46/348, *A.P.U. Report No.* 1.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quart. J. exp. Psychol.* **1**, 6–21.

Mackworth, N. H. (1950). *Researches on the Measurement of Human Performance.* No. 268, Spec. Rep. Series., Med. Res. Council, H.M. Stationery Office, Kingsway, London.

Pepler, R. D. (1953). The effect of climatic factors on the performance of skilled tasks by young European men living in the tropics. 4. A task of prolonged visual vigilance. *A.P.U. Report No.* 156/53.

Ross, B. M. (1955). Randomization of a binary series. *Amer. J. Psychol.* **68**, 136–8.

Williams, S. B., Bartlett, N. R. & King, E. (1948). Visibility on cathode-ray tube screens: screen brightness. *J. Psychol.* **25**, 455–66.

Woodrow, H. (1951). Time perception. In Stevens, S. S. (ed.)., *Handbook of Experimental Psychology*. N.Y.: Wiley.

(*Manuscript received* 9 *July* 1957)

INCREASE IN PERCEPTUAL SENSITIVITY AS A FUNCTION OF LEARNING IN THE TEST SITUATION

By J. D. KEEHN

American University of Beirut

An experiment in which subjects showed either significantly heightened or significantly lowered recognition thresholds of words previously failed in an anagrams test has been interpreted as demonstrating 'individual differences in reactions to threats of self-esteem'. A more specific explanation in terms of learning during the experiment is offered. It is argued that if anagrams constructed from highly familiar words are failed (due to insufficient time) no further learning can occur when the anagram solution is given, and hence no differences in recognition thresholds as between failure and success words will be found. Experimental findings supporting this contention are reported.

I. Introduction

The concept of perceptual defence was introduced as a way of describing an experimental finding that recognition thresholds of words in 'low value areas' were higher than positive values or neutral words (Postman, Bruner & McGinnies, 1948). Although the concept has since been rejected consistently by at least one of its authors (Postman, 1953), its cause has been taken up by other writers, particularly Eriksen (1954*a*).

One of the experiments given prominence by Eriksen in support of the 'perceptual defence' concept is that of Postman & Solomon (1950). In this study it was demonstrated that failure to solve anagrams in an apparently competitive situation resulted in subjects showing either significantly greater or significantly less difficulty of recognition of failure words, as compared with success words, in a subsequent brief exposure test. Although Eriksen accounted for these findings in terms of 'individual differences in reaction to threats of self-esteem' no precise account of how this comes about was attempted. He did, however, try to relate the general perceptual defence findings to those showing the effects of anxiety on learning. Eriksen treated the recognition of briefly exposed stimulus words as a problem-solving situation, and by assuming that anxiety interfered with the ability to recognize and report words from fragmentary cues argued that 'stimuli provoking anxiety may require more cues before correct recognition occurs' (Eriksen, 1954*a*). Apart from the difficulty of requiring that words must be 'recognized' as anxiety provokers before they are perceived and correctly reported (a difficulty which Postman (1953) has tried to overcome), the fact that some words which might be expected to provoke anxiety are recognized *sooner* than other words is not easily explained in this way.

The purpose of the present report is to attempt to account for some of the findings commonly represented as 'perceptual defence' using the Postman & Solomon (1950) experiment as a model.

II. Rationale

In the Postman & Solomon study ten anagrams were constructed from words which, according to the Thorndike–Lorge word count, were 'moderately unfamiliar'. After subjects had solved or failed to solve these words as anagrams they were subsequently tested for recognition thresholds on the same words with the results described above.

Although all the words which they used may have been of about the same familiarity level in general, the fact that they were all 'moderately unfamiliar' leaves open the possibility that they may have represented different degrees of familiarity to the individual subjects. Clearly we should expect subjects to find it more difficult to unscramble letters forming words they do not know well than to unscramble those forming familiar words. As it is also generally recognized that within limits familiar words have lower recognition thresholds than unfamiliar ones (Solomon & Howes, 1951), the fact that some subjects recognized success words sooner than failure words needs no mediating 'defence' or other hypothesis. Unfamiliar words take longer to recognize and are more difficult to reconstruct from anagrams, other things being equal.

Thus those subjects in the Postman & Solomon study taking significantly longer to recognize the anagram words which they failed to solve are of little theoretical importance in this case. It is the instances of heightened sensitivity to the failed words which are of interest.

If we regard the first part of the Postman & Solomon study as a kind of learning experiment (since after each anagram subjects were told the correct solution), the subsequent higher sensitivity of some subjects to failed words can be accounted for in terms of individual differences in learning the failed words. Conversely, subjects failing to benefit from the learning situation would show less sensitivity, owing to continued unfamiliarity.

If heightened sensitivity arises as a result of learning, then, if highly familiar words are used in the first place, so that no room is left for improvement within the experimental situation, clearly no differences would be expected between recognition thresholds for success or failure words. On the other hand, if we are dealing with some independent concept such as 'reaction to threat' it would appear reasonable to suppose that the simpler the words failed the greater the perceived threat, and thus the higher the subsequent recognition threshold for those words. Here, then, we would seem to have a crucial test. If individual differences in 'perceptual defence' are accountable for by something like the immediately prior experience of learning, repetition of the Postman & Solomon (1950) experiment with familiar words should result in no significant differences between thresholds for success and failed words. If, however, 'perceptual defence' is a function of here-and-now avoidance of threat, the differences found by Postman & Solomon should be amplified owing to the increased threat from failing in a simple situation.

III. PROCEDURE

Method. Fifteen six-letter words were used, ten easy and five difficult. The easy words constituted the experiment proper, the difficult words acting as controls to see whether the Postman & Solomon results could be replicated with them. The easy words were STRONG, CLOUDY, POLITE, PENCIL, LEADER (or DEALER), ANSWER, NEARLY, BARBER, STREAM (or MASTER), and ORANGE; the difficult words were BELFRY, SEPTIC, STANZA, PEWTER and MINGLE.* The difficult words were mainly selected from the Wechsler vocabulary tests and were hardly ever defined correctly by subjects with similar backgrounds to those used in this experiment. The easy words were similarly hardly ever failed when used in a vocabulary test independently of the present study.

---

* The word *mingle* is not in the Wechsler vocabulary lists; it was included because only four suitable six-letter words were available from Wechsler's tests.

The above fifteen words were presented in booklets designated as intelligence tests in the following anagram forms respectively: NTSROG, LCODUY, OPTLIE, PNCIEL, ERDELA, EWANSR, LYRNEA, RBBAER, STERAM, NOARGE, YLBEFR, ESPCIT, ZNTSAA, WTPEER and GMNIEL. Each anagram was printed on a separate sheet and 1 minute was allowed for each solution; this short time was used to ensure that some of the easy words would be failed. Subjects were instructed to raise their hands on the completion of each solution, the experimenter whispering audibly 'good' occasionally to give the impression that many subjects were finishing even when few such indications were made. After each anagram the correct answer was given and written by the subjects.

At the completion of the anagram section subjects were given 5 minutes to write as many English words as they could beginning with the letter '*t*'. This procedure was taken over from Eriksen (1954*b*) and was intended to supply a uniform interim period between the first and the second parts of the experiment.

Finally the fifteen words were presented tachistoscopically in random order. With the apparatus available, a Keystone Flashmeter attached to a Keystone Overhead Projector, model 1055, it was impossible to replicate the Postman & Solomon (1950) procedure exactly. Instead, 20 exposures were given, 3 at 1/100 sec., 5 each at 1/50, 1/25 and 1/10 sec. and 2 at 1/5 sec. The mode and rationale of scoring followed that of the earlier study. Although this scoring procedure is slightly coarser than that of Postman & Solomon and could possibly lead to attenuation of their findings, the comparison of easy and difficult words allows a check on this.

Subjects. In all, thirty subjects from an introductory course in Psychology were tested, in three groups of ten each. The separate experimental sessions lasted about an hour each, of which half was taken up with the tachistoscope test. A preliminary trial was run on ten subjects prior to the experiment proper to ensure a smooth testing procedure. Results from these subjects are not included in the final analysis.

IV. RESULTS AND DISCUSSION

Altogether the whole group of thirty subjects obtained 27·67 % failures on the easy and 81·29 % failures on the difficult anagrams. These figures are neither surprising nor important except to confirm our choice of easy and difficult words for making the anagrams. The essential comparison is that between speed of recognition and anagram success or failure for the individual subjects on the easy words. It is this comparison that Postman & Solomon (1950) found important with moderately unfamiliar words, and in our case the only apparent relevant modification of their procedure is the use of highly familiar as against moderately unfamiliar words.

For the statistical purposes of this analysis only those subjects who made three or more errors on the easy anagrams were retained. Although this number is extremely small for effecting comparisons, it is comparable to the numbers in the Postman & Solomon study. Sixteen subjects met this criterion; all of them failed also to solve three or more of the difficult anagrams. Column 't_1' in Table 1 shows that only one of these subjects differed significantly in the number of exposures required for success and failure words in the easy list, as against eleven of seventeen subjects showing such differences in the Postman & Solomon (1950) experiment.

Table 1. *Mean number of brief exposures required to recognize easy errors, easy successes and difficult errors respectively by individual subjects*

(t_1: comparison of easy errors against easy successes. t_2: comparison of easy successes against difficult errors.)

Subject	Easy errors	Mean exposures	Easy successes	Mean exposures	t_1	Difficult errors	Mean exposures	t_2
1	4	2·00	6	2·67	N.S.	5	4·0	N.S.
2	4	4·75	6	5·33	N.S.	5	11·8	0·05
3	5	6·00	5	2·80	N.S.	4	11·5	0·05
4	4	12·50	6	9·67	N.S.	5	16·2	N.S.
5	3	11·33	7	8·57	N.S.	4	15·0	0·02
6	3	2·67	6	3·33	N.S.	5	3·8	N.S.
7	3	6·00	6	3·33	N.S.	5	10·2	0·05
8	3	2·33	6	2·17	N.S.	5	7·2	0·01
9	4	2·25	6	2·50	N.S.	5	8·8	0·05
10	4	3·25	6	2·17	N.S.	4	3·25	N.S.
11	4	6·75	6	4·50	N.S.	5	13·00	0·02
12	4	3·50	6	9·00	0·05	3	14·00	N.S.
13	3	5·67	7	4·29	N.S.	4	12·5	0·02
14	5	4·20	5	2·80	N.S.	5	5·8	N.S.
15	3	2·00	7	2·86	N.S.	4	6·6	0·02
16	4	7·50	6	5·50	N.S.	5	8·8	N.S.

Thus our contention that increased sensitivity to briefly exposed words can be attributed to some subjects having learned the words they had previously failed is supported, for when learning is not likely increased sensitivity does not occur.

Control data are presented in column 't_2' in Table 1. This column shows the comparison between number of exposures required to recognize easy words which were solved as anagrams, and difficult words which were failed. In nine of the sixteen cases a significantly longer time was taken to recognize the failed, or difficult, words.

This finding replicates that of Postman & Solomon (1950), except that all the differences were in the direction of showing failure to be related to lowered sensitivity. However, in this case this is to be expected since only difficult failed words are being considered.

Thus our re-interpretation of the Postman & Solomon findings is supported by the present experiment.

This study was made possible by a grant from the Rockefeller Brothers Fund to the American University of Beirut. The author wishes to acknowledge his indebtedness to the Fund and to Prof. Wayne Dennis for a critical reading of a draft of the manuscript.

REFERENCES

ERIKSEN, C. W. (1954*a*). The case for perceptual defense. *Psychol. Rev.* **61**, 175–82.

ERIKSEN, C. W. (1954*b*). Psychological defenses and 'ego strength' in the recall of completed and incompleted tasks. *J. Abnorm. Soc. Psychol.* **49**, 45–50.

POSTMAN, L. (1953). On the problem of perceptual defense. *Psychol. Rev.* **60**, 298–306.

POSTMAN, L., BRUNER, J. S. & MCGINNIES, E. M. (1948). Personal values as selective factors in perception. *J. Abnorm. Soc. Psychol.* **43**, 142–54.

POSTMAN, L. & SOLOMON, R. L. (1950). Perceptual sensitivity and incomplete tasks. *J. Personality*, **18**, 347–58.

SOLOMON, R. L. & HOWES, D. H. (1951). Word frequency, personal values and visual duration thresholds. *Psychol. Rev.* **58**, 256–71.

(*Manuscript received* 12 *November* 1957)

FIGURAL AFTER-EFFECTS, RETINAL SIZE AND APPARENT SIZE

By PETER McEWEN

Psychology Department, Queen's University, Belfast

Previous work has established that figural after-effects are determined by the retinal size of the inspection figure, but Sutherland (1954) has argued that the methods used were based on the assumption that retinal and apparent sizes cannot both be determinants. He himself has provided evidence of the effects of apparent size. The present paper reports an experiment intended to confirm Sutherland's findings and to measure the magnitude of the after-effects, if any. Eleven subjects were required to equate two test-figures after prolonged fixation of prior inspection figures. One inspection figure was retinally equal, apparently larger, the other, retinally equal, apparently smaller than a test-figure. In the first case, the test-figure decreased significantly, in the second it decreased, but not significantly. It is held that these results can be explained by assuming both retinal and apparent sizes to be effective, in the first case acting together and in the second opposing one another.

I. Introduction

Figural after-effects are perceptual distortions in test-figures following moderately prolonged stimulation, as, for example, from fixation of a visually presented form. The phenomena are by now sufficiently well known from the writings of, among others, Gibson, Osgood and Heyer and, particularly, Köhler and his collaborators. The work as a whole has been reviewed elsewhere by the present author (McEwen, 1958). The problem to be treated here originated with Prentice (1947), who was concerned with the effect on the magnitude of figural after-effects of varying the distance from the observer. The experienced size of an after-image depends upon the size of the area of the retina which is differentially adapted and on the distance of the surface on which the image is projected. Its size varies approximately linearly with the distance of the surface (Emmert's Law). Eidetic images, on the other hand, are said not to conform to this law, but these are considered to be cortical in nature and influenced by data other than visual data. They are difficult to study while figural after-effects, which are also said to be cortical, are relatively more manageable. Prentice raised the question whether figural after-effects have an absolute size, governed only by the size of the 'satiated' area, or whether their size is relative, as is the size of the after-image.

Prentice had as an inspection-figure a rectangle placed above and to the right of the subsequently presented test-figure. The test-figures were squares, on opposite sides of a fixation mark and horizontally aligned. The expected after-effect was a downward displacement of the right test-figure, measured by having the subject re-align the left test-figure with it. The subject inspected at 2 m. and had the test-figures presented at 2 and 6 m.

Prentice argued in this way: the size of the satiated area remains the same throughout, while test-figures of the same objective size are viewed at different distances. The apparent size of the latter will remain very much the same throughout since there are cues permitting a high degree of constancy. Therefore, if apparent size is the effective factor, there will be no difference in the measured size of the after-effect at 2 and at 6 m., although angular size will differ. But if the satiated area is 'projected' on a more distant background and covers a larger area of it, the displacement of the test-figure will be greater since the contour of the test-figure will be closer to the satiated region. He found no differences. But since at 6 m. the retinal size of the test-figure is smaller than at 2 m. and its contour is three times farther from the inspection figure, and since this decrease could cancel out any increase due to reduced angular separation, Prentice repeated the experiment with other distances. When again he found no differences in after-effects, he concluded that apparent size determined the magnitude of after-effects.

In his next paper Prentice (1950) explains his reasons for this conclusion in more detail: 'the data of our two experiments show that *no change* occurs in the size of the measured figural after-effect whether comparisons are made at 3, 5 and 7 m. or at 2 and 6 m.; this fact implies the shrinking of the retinal image *exactly* cancels the increase due to visual angle at several distances; but that increase is known to be precisely linear and therefore the function relating distance of test-contour from satiated region to size of distortion would also have to be exactly linear to account for the "balancing out" at any distance; the latter function is known to pass through a maximum (Köhler & Wallach) and there is no evidence to indicate that it becomes straight after that; the coincidence seems highly improbable.' In this same paper Prentice reports findings which cast doubt on his first conclusion. Thouless, in correspondence with Köhler, had reported an unpublished experiment which involved pitting retinal and apparent sizes against one another in circumstances in which the decision would rest upon the direction, rather than the magnitude, of the after-effects. Adopting this method, Prentice had his subjects view an inspection circle which was retinally smaller but apparently larger than a test-circle, placed nearer to the subjects. In this case, if it is the retinal size which determines the after-effect, the affected test-circle should look larger than its comparison test-figure but if apparent size is operative, the test-figure should look smaller. Of sixty-five subjects, forty-five (or 69%) reported the affected test-circle as being increased in size, a finding which favours the interpretation that retinal size determines the after-effect.

Hochberg & Bitterman (1951) improved on Prentice's work in that they measured the displacements. In their experiment the retinal size of the inspection figure was larger than that of the test-figure, the apparent size smaller. Their findings supported those of Prentice's 1950 experiment.

At this point the matter seemed almost settled but, although Prentice (1950) rejects the reasoning which had made him arrive at the opposite conclusion in his earlier paper, its results remain unexplained. Osgood & Heyer (1952), writing apparently without knowledge of the later investigation, and satisfied that a conclusion in favour of apparent size would be at variance with their statistical theory, considered that Prentice's earlier results must be questioned on two grounds: first, that he had ignored the distance paradox—this, his second paper makes plain, is not the case—and secondly, that physiological nystagmus will result in greater variability of judgement at greater distances. There is no suggestion in Prentice's account that such increased variability was found.

An alternative proposal is made by Sutherland (1954), who points out that all previous investigations assume that retinal and apparent size are mutually exclusive as determinants of after-effects. But if Prentice's second investigation and that of Hochberg and Bitterman are accepted as demonstrating the effect of retinal size under some conditions, and Prentice's first investigation is taken to show that apparent size is also a determinant the inconsistency in reported results disappears. Both retinal and apparent sizes are effective, and experiments which pit them against one another prove only that there are conditions where the effect of one or other may be concealed.

Sutherland performed three experiments. In the first, two inspection and test-circles were placed at different distances and were of such sizes as to subtend the same angle at the retina, but were of different apparent sizes. (Retinal sizes were the same in order to render retinal size ineffective. According to the theoretical accounts proposed by Köhler & Wallach (1944) and Osgood & Heyer (1952), it would be predicted that where the contours of inspection and test figures coincided no displacement would occur.) In the first experiment, the inspection figure had the larger apparent size and subjects saw the affected test-figure as smaller (19, as against 5 who saw it larger, a difference which was significant). In the second experiment, the inspection figure had the apparently smaller size, and nineteen out of twenty-two subjects saw the test-circle as larger than its partner.

In Sutherland's third experiment the retinal size of the inspection circle was larger, the apparent size smaller than that of the test-circle, the arrangement used by Hochberg & Bitterman. The results were in agreement with Hochberg & Bitterman's finding.

Sutherland notes that the result of his first experiment could be explained without reference to apparent size, since Köhler and Wallach had, despite their theoretical account, found that when inspection and test-contours coincide, the test-figure appears to shrink slightly. This could account for his subjects seeing the test-figure as smaller in the first experiment, but the results of the second experiment would remain unexplained. He concluded that his experiments demonstrate that both retinal and apparent size may create after-effects and that both may be effective in producing a given after-effect.

II. THE PRESENT INVESTIGATION

Sutherland's finding complicates any attempt at a theoretical account of figural after-effects and it seemed desirable not only that it should be confirmed, but that the magnitude of the effects should be established in the way in which Hochberg & Bitterman did for Prentice's (1950) finding. This was the objective of the present study.

(i) *Apparatus*

The problem required two adjustable test-figures, one being the affected figure, the other for comparison and to be adjusted to look equal in size to the affected one; two inspection-figures, equal in retinal size to the affected test-figure when placed at appropriate distances from the subject, but one to be apparently larger, the other apparently smaller than the test-figure. It was also necessary to make observations with an inspection-figure in one other position, so that both retinal and apparent sizes were different from that of the test-circle, in order to establish that after-effects occurred predictably with the particular apparatus and subjects since, in the main experiment, negative results might be found.

The inspection and test-figures were set in three wooden screens painted a flat black. Screen TT, 24 by 15 in., contained the test-figures and a fixation point. The test-figures were two illuminated disks to left and right of the fixation mark and on a level with it horizontally. The fixation mark was an illuminated red spot in the centre of the screen. To make the test-figures, two circular holes were cut in the screen, in each of which an iris diaphragm was set. Between the leaves and the casing of the diaphragm, a disk of parchment was inserted, in such a way that, except when the diaphragm was fully open, only a white disk of parchment could be seen. These disks, when illuminated from behind by a lamp in a canister attached to the back of the screen over the diaphragm, were found to provide the most satisfactory way of obtaining lighted circular test-figures of variable diameter, without refraction. The levers controlling the diaphragm aperture protruded through a slot in each canister and were extended to provide pointers reading against large protractors.

Each of the other two screens had a similar fixation mark and an inspection-figure to the left as viewed by the subject. Each figure was illuminated from behind in the same way as the test-figures. Screen S. 24 by 15 in., had a circular hole 15 mm. in diameter, the centre of which was 5 cm. from the fixation mark. In this was set a disk of opal glass, painted black but for a small ring with outer diameter 15 mm.. and inner diameter 11 mm. This ring, when illuminated, was the smaller inspection-figure. Screen L was 48 by 30 in. with a similar illuminated ring as the larger inspection-figure. The centre of this ring was 20 cm. from the fixation point; its outer diameter was 64 mm. and its inner diameter 60 mm. Since this screen was to be at a greater distance from the subject, it was made larger than the others to avoid possible after-effects arising from its contours.

Throughout the experiment, the fixation marks were at eye-level. The test-screen was directly in front of the subject, 150 cm. from his eye. The standard disk, the affected one, was set at 30 mm. diameter. The other screens were used alternately. Screen S was 45° to the subject's right, 75 cm. from his eye and, when viewed directly, the outer diameter of the inspection-figure (15 mm.) subtended the same angle at the eye as the standard disk but appeared to be smaller. Screen L, with a figure of inner diameter 60 mm. was placed 35° to the subject's right, 300 cm. from his eye. At this distance, this ring also subtended the same retinal angle as the standard disk but appeared to be larger. Screen L was also used in another position, 250 cm. from the subject and 35° to his right, with the fixation mark moved appropriately. In this position, the inspection-figure was both retinally and apparently larger than the standard disk.

During the experiment the room lighting was low, giving only dim illumination, but clear outlines to the figures, which were more brightly lit. The light-grey background walls gave sufficiently strong cues to the relative distances of the screens. Subjects were asked to confirm that the screens were in fact seen to be at different distances. Joynson (1949) has pointed out that a normal perceptual attitude gives constancy where the distances are within 120 ft., where binocular vision is used, where the objects are placed in different directions from the eye, and where unusual external conditions such as reduction screens, reduced illumination, etc., are avoided. These conditions, except the fourth, are met here, but although the lighting was reduced the lighter background helped to give good perception of relative distances. The subjects were not asked to judge sizes at different distances; the test-figures were at the same distance, and, to the subjects, the inspection-figures were 'the large circle' and 'the small circle' on the screens fixated before making judgements. It is hardly likely that any analytic attitude towards relative sizes and distances was present to reduce the ordinary perception of constancy.

(ii) *Subjects*

The subjects were eleven first-course students of psychology, eight men and three women, naïve as regards the phenomena of figural after-effects, and all with normal or corrected vision.

(iii) *Procedure*

Before each experimental session, subjects were partially dark-adapted. The subject was seated at a table, and instructed in the use of the head-rest and on the importance of fixating throughout the main part of the experiment. In the first two sessions the test screen alone was used and the subject practised making the variable disk equal to the standard one. A modified method of limits was used, the experimenter altering the variable from alternately 'larger' and 'smaller' starting positions on instructions from the subject. After each trial in the first and for half the second session the subject was given information about his performance, but after this he worked without knowledge. In each session two sets of ten trials were given with a rest between sets. In the third session, experimental trials began. Each session began with ten trials without prior fixation of the inspection figure. These ten trials formed a control set for the following experimental condition. (This procedure resembles that used by Hammer (1949), whose data are among the best yet obtained on figural after-effects.) Following the control trials in each session, one of three conditions was used. The order of presentation of the conditions was systematically varied, and at least one day separated successive sessions.

(1) *Condition L*

S fixated the screen with retinally equal but apparently larger inspection-figure for 2 min., transferred his gaze to the test-screen and adjusted the setting of the test variable to apparent equality with the standard.

(2) *Condition S*

Fixation on screen S, with retinally equal but apparently smaller inspection-figure; times as above.

(3) *Condition X*

S fixated the larger screen, L, now placed at 250 cm. so that the inspection-figure was both retinally and apparently larger; times as above. Only six of the eleven subjects served in this condition.

(iv) *Results*

The results are shown in Table 1. For each individual subject, a mean control and a mean experimental setting was found for each of the three conditions. Then for each condition, the difference between each subject's control and experimental means was found, these differences were summed over the group and a mean difference, a standard error of the mean difference and a value of *t* were obtained. These are shown in the table. Negative values indicate that the affected test-circle decreased in size.

Table 1 *Group mean displacements for the three conditions*

Condition	Group mean difference (mm.)	S.E. of mean difference	D.F.	*t*
X	−1·32	0·352	5	3·75*
L	−1·01	0·310	10	3·26†
S	−0·23	0·220	10	1·05

* indicates 5% significance. † indicates 1% significance.

BUREAU RESEARCH LIBRARY

III. DISCUSSION

In condition X, where the inspection-figure is both retinally and apparently larger than the test-figure, for all subjects the test-figure decreases in size and the mean decrease is significant, so it may be taken that figural after-effects occur with this apparatus and these subjects in a predictable way.

With condition L, the inspection-figure has the same retinal size as the test-figure and, according to both Köhler & Wallach and Osgood & Heyer, we should not expect any displacement, if retinal size *only* is effective. If apparent size is effective, we should expect a decrease in the size of the test-figure since the apparent size of the inspection-figure, in this condition, is larger than the size of the test-figure. In fact, the test-figure does shrink for all but one subject and the mean displacement is a decrease, significant at just below the 1% level.

With condition S, the retinal sizes of inspection and test-figures are the same and again we should expect no displacement, unless the smaller apparent size of the inspection-figure is effective, in which case an increase in size of test-figure would be expected. For six of the eleven subjects, the test-figure decreases; for the remainder it increases and for the group, the mean change is a slight and insignificant decrease.

At first sight, the results suggest contradictory interpretations. Where the retinal size of the inspection-figure is the same as that of the test-figure and the apparent size is greater, the test-circle shrinks, suggesting that the greater apparent size is the decisive factor. But where the apparent size of the inspection-figure is smaller, while no significant after-effect is found, the displacement is in the opposite direction to that predicted on the assumption that apparent size is effective. But, while Köhler & Wallach and Osgood & Heyer claim that where the contours of inspection and test-figures coincide there will be no displacement, Köhler & Wallach report observations in which the test-figures decreased in size, in the cases of both circular and rectangular figures. They also state that a certain subject, who had had the largest figural after-effects in quantitative studies, was unable to detect more than a just perceptible *shrinkage* when the inspection and test-figures coincided in space. Walthall (1946) reports observations which agree with those of Köhler & Wallach. More conclusive than these, however, are the data reported by Sagara & Oyama (1957) in a recent review of Japanese studies. According to these writers, no fewer than six Japanese experiments have studied the magnitude of displacement as a function of the size ratio between inspection and test-circles (Ikeda, 1951; Oyama, 1952; Ikeda & Obonai, 1955; Kogiso, unpublished). These studies agree in finding that where inspection and test-circles coincide the test-figure shrinks, five of them agreeing that the decrease in size is about 5%.

We may now argue as follows:

(1) If retinal size alone is effective, then, in both conditions L and S, we should expect to find much the same displacement in the same direction—towards a decrease in test-size. But while both are in the same direction, only the L displacement is significantly different from zero. It is also significantly different from the S displacement at below the 1% level.

(2) If apparent size alone is effective, we should expect to find displacements in opposite directions, both, presumably, significant. Only the L displacement is significant and the direction is the same in both conditions.

(3) If both retinal and apparent sizes are effective, then in condition L, where both are expected to act in the same direction, a significant decrease in the size of the test-figure should be, and is, found. But in condition S, the two factors are expected to counteract one another and any displacement occurring must be reduced, the observed direction being unpredictable. In condition S an insignificant decrease is found.

It would seem that the interpretation which best fits the data is that both apparent and retinal sizes determine the after-effect. On the ground that the displacements are in the same direction, it could be maintained that the retinal size alone would produce the observed effects, but then the significant difference in magnitude of the after-effects between conditions L and S would remain to be explained. It is concluded that both retinal and apparent sizes are operative.

It should be mentioned that in the review by Sagara & Oyama it is said that Oyama (1954, 1956) has obtained results in which it was clear that retinal size was a determinant and that it was possible that apparent size was also, though the results were not conclusive.

The findings reported call for a more complex theory of figural after-effects than any so far put forward. As Sutherland has pointed out, they mean that perception is at least a two-stage process, and that provision will have to be made in theory for the action of the process determining apparent size on that determining after-effects. It has been pointed out elsewhere by the present writer (McEwen, 1958) that some of the problems of a theoretical account would be made less difficult if it were shown that part, at least, of the process, which is observed as a figural after-effect, took place at a retinal level. The results of the present experiment would seem to make this more likely, for if there are two stages in the development of the after-effect, the retina seems the most likely, locus for the first stage. In this connexion, also, the paper by Motokawa, Nakagawa & Kohata

(1957) should be noted. These investigators (using Motokawa's 'retinal induction' technique) believe that they have demonstrated that displacement and the distance paradox can be explained by retinal induction.

REFERENCES

HAMMER, E. R. (1949). Temporal factors in figural after-effects. *Amer. J. Psychol.* **62**, 337–54.
HOCHBERG, J. E. & BITTERMAN, M. E. (1951). Figural after-effects as a function of the retinal size of the inspection figure. *Amer. J. Psychol.* **64**, 99–102.
IKEDA, H. (1951). Studies in figural after-effects. Tokyo Bunrika University Dissertation. (Unpublished.)
IKEDA, H. & OBONAI, T. (1955). The studies in figural after-effects. IV. The contrast-confluence illusion of concentric circles and the figural after-effect. *Jap. Psychol. Res.* **2**, 17–23.
JOYNSON, R. B. (1949). The problem of size and distance. *Quart. J. Exp. Psychol.* **1**, 119–35.
KOGISO, I. Unpublished data referred to in Sagara & Oyama.
KÖHLER, W. & WALLACH, H. (1944). Figural after-effects. *Proc. Amer. Phil. Soc.* **88**, no. **4**, 269–357.
MCEWEN, P. (1958). Figural after-effects. *Brit. J. Psychol.* (Monogr. Suppl.).
MOTOKAWA, K., NAKAGAWA, D. & KOHATA, T. (1957). Figural after-effects and retinal induction. *J. Gen. Psychol.* **57**, 121–35.
OSGOOD, C. E. & HEYER, A. W. Jr. (1952). A new interpretation of figural after-effects. *Psychol. Rev.* **59**, 98–118.
OYAMA, T. (1952). Figural after-effects and reversible figures. *Chiwa Sensei Kanreki Kinen Ronbunshu*, pp. 47–55.
OYAMA, T. (1954). Experimental studies of figural after-effects. II. Spatial factors. *Jap. J. Psychol.* **25**, 195–206.
OYAMA, T. (1956). Temporal and spatial factors in figural after-effects. *Jap. Psychol. Res.* **3**, 25–36.
PRENTICE, W. C. H. (1947). The relation of distance to the apparent size of figural after-effects. *Amer. J. Psychol.* **60**, 617–23.
PRENTICE, W. C. H. (1950). On the size of the figural after-effect with varying distances. *Amer. J. Psychol.* **63**, 589–93.
SAGARA, M. & OYAMA, T. (1957). Experimental studies on figural after-effects in Japan. *Psychol. Bull.* **54**, 327–38.
SUTHERLAND, N. S. (1954). Figural after-effects, retinal size and apparent size. *Quart. J. Exp. Psychol.* **6**, 35–44.
WALTHALL, B. J. (1946). The Köhler effect. *Amer. J. Psychol.* **59**, 152–5.

(*Manuscript received* 10 *December* 1957)

EYSENCK'S THEORIES OF ANXIETY AND HYSTERIA—A METHODOLOGICAL CRITIQUE

By VERNON HAMILTON

Institute of Psychiatry, University of London[*]

An examination is made of the evidence for and the validity of Eysenck's dimension of introversion-extraversion. The dimension is found to be deficient on both counts, and therefore considered unsuitable as a starting-point for further hypotheses. Eysenck's postulate and hypotheses linking introversion-extraversion and anxiety and hysteria with cortical and reactive inhibition and with conditionability, satiation phenomena, reminiscence effects, etc., are discussed in relation to their predictive value and to the sufficiency of the evidence which has led Eysenck to conclude that his hypotheses have been confirmed.

On the basis of a detailed review of Eysenck's findings, it is here concluded that the relationship between the continua of anxiety-hysteria and of introversion-extraversion has not been sufficiently demonstrated, and that the relationship between either of these continua and reactive inhibition and cortical inhibition represented by conditionability and satiation effects has not been established. It is further contended that there is insufficient empirical evidence for the identity of the inhibitory concepts used in the theoretical framework.

I. INTRODUCTION

During the last decade Eysenck and his colleagues and students at the Institute of Psychiatry have undertaken and published a vast amount of research aimed at establishing a scientific basis for the study of personality. The theories and views put forward have always been provocative and almost always have stimulated new lines of thought. For an appraisal of the theories of social behaviour, opinions and political views, the reader is referred to the articles by Rokeach & Hanley (1956), Hanley & Rokeach (1956) and Christie (1956*a*, *b*). It is also not within the scope of the present review to re-enter the controversy concerning the status of conclusions based on factor analysis and criterion analysis in relation to the requirements of a hypothetico-deductive system (Albino, 1953; Eysenck, 1953*b*).

The newly developed causative, dynamic theory of anxiety and hysteria put forward first by Eysenck (1955*a*) with confidence, and latterly (Eysenck, 1957*b*) with caution, would seem to rest primarily on the evidence for an introversion-extraversion dimension and on the important experiments on conditioning and the figural after-effect. For that reason the task of the present critique will be limited to the following two objectives: (1) to examine the validity of Eysenck's dimension of introversion-extraversion and the evidence as to the validity of regarding its extreme variants—dysthymics and hysterics—as criterion groups for the dimension; and (2) to examine the value of hypotheses and findings that have been generated by assuming that the concepts of reactive and cortical inhibition and satiation are identical and provide causative principles for the explanation of the introversion-extraversion dimension.

The challenge extended by Eysenck (1955*a*, 1957*b*) that the critic of his theory must put forward an alternative and equally comprehensive theory, is not met here. It is held, instead, that the lack of evidence for most of the hypotheses and the inconclusiveness of other findings, as far as they have been presented to date, strongly suggests multiple causation where the various parameters need not be mutually exclusive.

---

[*] Now at Springfield Hospital, Upper Tooting, London, S.W.17.

II. THE EVIDENCE FOR THE DIMENSION OF INTROVERSION-EXTRAVERSION

Eysenck's theory of the relationship between the dimension of introversion-extraversion and the current nosological categories of the neuroses is primarily based not on his own earliest investigations (1947), but on the findings of Hildebrand (1953). Since this study and the conclusions derived from it form one of the main assumptions underlying later postulates linking introversion-extraversion with individual differences in reactive inhibition, it obviously merits careful appraisal, especially as regards the validity of its conclusions. Eysenck (1955*a*, *b*) writes that the investigations of Hildebrand (1953) '...allow us to use the hysteric/psychopath group on the one hand and the dysthymic group on the other as criteria for any predictions made in terms of a theory of extraversion-introversion'. Elsewhere (1957*a*), Eysenck considers that the formal proof of the existence of the introversion-extraversion dimension '...has been accomplished in terms of objective tests which can be used as an operational definition of this dimension'. It is here proposed to examine critically the evidence on which these two conclusions are based.

Hildebrand (1953) administered a battery of three intelligence tests, six personality inventories, tests of static ataxia and suggestibility, the Porteus Mazes, an annoyance questionnaire, a humour scale and a leg persistence test to patients diagnosed as anxiety states, depressives, obsessionals, hysterics and psychopaths, and to normal control subjects. In the resultant correlation matrix, 18% of the correlations were significant. Three factors were extracted from the matrix: intelligence, neuroticism, and introversion-extraversion. Only four tests made a sizeable contribution to the introversion-extraversion factor: Guilford's Rhathymia and Social Introversion scales (R and S), the Porteus Mazes test, and the Maudsley Medical Questionnaire, accounted for a variance of 6% out of a total variance of 32%. Of these four, the Rhathymia scale most clearly defined the introversion-extraversion factor.

Table 1. *Extract of correlations and factor loadings from Hildebrand* (1953)

Correlations between R, S, and M.M.Q. scales

	R	S	M.M.Q.
R	—	−0·487	−0·305
S		—	0·644
M.M.Q.			—

Factor loadings

	R	S	M.M.Q.	Porteus W.W.D.	Porteus C.L.
Factor I: neuroticism	−0·25	0·74	0·74	0·26	0·07
Factor II: introversion-extraversion	0·55	−0·47	−0·26	0·36	0·43

(W.W.D. and C.L. are weighted wrong direction and crossed lines scores.)

Table 1 shows that the S scale and the Maudsley Medical Questionnaire have their major loadings on the neuroticism factor, and that the major measures of introversion-extraversion and of neuroticism are significantly correlated. These results make it clear that the Guilford scales cannot produce an undiluted measure of either neuroticism or of introversion-extraversion and that orthogonal solutions of such a matrix are factor analytical artifacts rather than evidence for the unidimensionality of Eysenck's measures. The significant data from the Porteus Mazes test should also be interpreted cautiously since this test is of very limited clinical diagnostic validity.

On the Rhathymia scale, the Porteus Mazes and some scoring categories of the body sway test of suggestibility, the control group in Hildebrand's study behaved about as extravertedly as a group of patients with an hysterical neurosis (Hildebrand, 1953). It will be evident that these data are by no means sufficient evidence for an introversion-extraversion dimension. Hildebrand himself writes '...the evidence in this case is not so convincing as for the intelligence and neuroticism factors...and further statistical techniques are clearly indicated'. It is conceivable that the variance of the introversion-extraversion factor was so small because the criterion groups (dysthymics, hysterics and psychopaths) were excluded from the factorial analysis. However, the relatively small number of tests differentiating significantly between dysthymics and hysterics, and the small number of measures that were found to have significant loadings on the extraversion factor, suggest that it is doubtful whether a factorial analysis of the data from all experimental groups would have yielded a much greater factor. Unfortunately, this additional analysis does not appear to have been done.

To increase the evidence, estimated factor scores based on regression coefficients were calculated (Hildebrand, 1953) and the experimental groups were compared for their positions on the two factors of neuroticism and introversion-extraversion. This analysis, in addition to providing information about the relative contribution of the tests to the introversion-extraversion factor, gave a vital piece of information which Eysenck (1955*a*, *b*) omits to report. It was found that in terms of the tests used in the investigation, the normal control group had the biggest estimated factor score on extraversion, i.e. the normal control group was more extraverted in terms of the tests used than neurotic extraverts (Hamilton, 1956*b*).

This finding was clearly contrary to hypothesis. It does not suffice to say, as Hildebrand does, that the results obtained from the normal group show '...them to be highly extraverted in terms of the normal population as a whole'. For firstly this group *is* Hildebrand's normal population, and secondly such argument is circular since an extravert is defined as a person who scores highly on tests supposedly measuring extraversion. Hildebrand's claim that '...the battery measures introversion-extraversion efficiently among a normal population equally well as in a neurotic population' is, therefore, hardly substantiated. It is of interest to note that Eysenck (1957*b*) continues to support the *ad hoc* argument that Hildebrand's control group was abnormally extraverted and refers to a second control group employed instead of the '...unsuitable...' first group (1957*b*, p. 29). It may be considered unfortunate that neither the mean R scale scores of the two control groups nor their position in the dimensional diagram (Eysenck, 1957*b*, p. 30) are given. Reference to Table 2 will indicate that the so-called abnormally high score of Hildebrand's original control group is probably no more than 1 S.D. above the mean for tested normal control groups (cf. Eysenck, 1955*b*). In addition, it is questionable to what extent the substitution of suitable for unsuitable test data is a scientifically valid procedure, unless double data for all experimental groups are obtained.

In terms of correlation coefficients, loadings on the introversion-extraversion factor, and inter-group discriminability, Guilford's Rhathymia scale is seen to be the best test measuring the postulated introversion-extraversion factor, so that subsequently it has been used as an external selection criterion by Eysenck. One might almost be tempted to say that introversion-extraversion *is* what is being measured by the R scale. It will, therefore, be useful to analyse a little more carefully various data bearing on the validity of this scale.

The Greek word Rhathymia means freedom from care. In Lovell's factor analysis of the Guilford scales (Eysenck, 1953*a*) the superimposed rotated axes indicate that Rhathymia is found in the same quadrant as objectivity, freedom from depression, self-confidence, sociability etc. One needs therefore to examine whether the R scale can differentiate reliably between normal control subjects and hysterics or between neurotic extraverts, e.g. hysterics and heterogeneous groups of neurotic introverts whom Eysenck calls dysthymics.

Table 2. *Summary of mean R scale scores*

	Hysterics		Dysthymics		Controls	
	$\bar{x}$	N	$\bar{x}$	N	$\bar{x}$	N
Investigation:						
Eysenck (1955*b*)	40	14	25	14	—	—
Franks (1956)	49·6	20	20·6	20	40·5	20
George (1954)	—	—	—	—	36·3	500
Hildebrand (1953)	36·6	25	26·3	45	45·7	45
Martin (1955)	22·3	17	28·2	20	33·8	27
Nichols (1956)	36·9	27	31·1	26	—	—

Several conclusions may be derived from the data of Table 2. In only one study (Franks, 1956) is the difference between hysterics and normal controls significant as well as in the predicted direction. The data found by Martin (1955) and Hildebrand (1953) show higher R scale scores for control than for hysteric groups, and the normal subjects in George's large-scale English study (1954) show mean scores on the R scale almost identical with those found by Hildebrand for hysterics. There would appear to be sufficient grounds, therefore, to reject the scores of Hildebrand's hysteric group as being unrepresentative of hysterics by reference to a normal control sample, a procedure which Eysenck has apparently considered valid with respect to Hildebrand's normal controls. As already indicated above, a comparison of Hildebrand's normal controls with other control subjects tested on the same scale does not support Eysenck's and Hildebrand's special plea that Hildebrand's normal control subjects were abnormally extraverted.

Table 2 also indicates that the dysthymic groups have fairly consistently lower means on the R scale than the other groups. The differences between hysterics and dysthymics on this scale was found to be significant on a one-tail test of significance and in the predicted direction in all but Martin's sample, but the overlap of all the data from the groups of controls, hysterics and dysthymics is considerable. It is thus clear that the R scale as a measure of extraversion and as an external criterion of clinically diagnosed hysteria produces a large error of misclassification. This would account for its inability to discriminate clearly between groups of dysthymics and hysterics rather than lack of large enough groups of experimental subjects as held by Eysenck (1955*b*).

Two other findings are relevant to the present discussion of the R scale as a valid measure of normal and neurotic extraversion. In Hildebrand's study the R scale discriminated significantly, and unexpectedly, between his normal and his neurotic groups, and it correlated to the extent of almost 0·5 with the Social Introversion scale of the Guilford battery. Since Hildebrand's normal group cannot be considered neurotic in terms of clinical assessments, of scores on the Maudsley Medical Questionnaire, nor by definition, it is possible that high scores on the R scale may well be due to lack of insight or to that old foe of questionnaire investigators—deliberate or unconscious falsification. The negative correlation between the S and R scales suggests, moreover, that *social* extraversion is perhaps the most important variable measured by the R scale. If that conclusion were correct, it would be hardly surprising if the scale did not differentiate reliably between hysterics and other groups since florid hysteria is only one of several clinical manifestations of hysteria. A great proportion of hysterics are in fact socially withdrawn, or inadequate, or both.

Within the general framework of Eysenck's theory of introversion-extraversion an additional and largely untested assumption is made concerning the homogeneity of the category of dysthymia and its subgroups of anxiety states, depressives and obsessionals. Evidence has been presented elsewhere (Hamilton, 1957*a*) that the concept of dysthymia lacks proven validity even for the strictly limited objective measures employed so far by Eysenck. It was argued that similarities between the subgroups of the dysthymia category may be a function of the test used and the mechanism that is being investigated, and that the homogeneity of dysthymia is not maintained when testing for functionally derived indices of behaviour such as conflict avoidance (Hamilton, 1956*a*, 1957*b*).

It would appear from the preceding discussion that the statistically best measure of introversion-extraversion is of insufficient validity and clinical discrimination to support the hypothesis of an introversion-extraversion dimension, particularly with regard to its extreme clinical varieties. In subsequent theorizing the dimension should, therefore, not be used for the construction of further hypotheses unless more conclusive evidence has first been obtained. Unless this is done, the later hypotheses would be at what Eysenck himself (1957*a*, *b*) calls the notional level of explanation and investigation. Thus any further deductions concerning conditionability and figural after-effects (Eysenck, 1955*a*) would start off from possibly false assumptions, so that the new hypotheses would have a high chance of being falsified on these grounds alone.

III. REACTIVE INHIBITION AND THE INTROVERSION-EXTRAVERSION DIMENSION

The following pages will be devoted to an appraisal of the theoretical background of the two major deductions made by Eysenck, and of the experiments to test them, on the basis of an attempted link between his theory of introversion-extraversion, Pavlov's theory of cortical excitation-inhibition and Hull's concept of reactive inhibition. Eysenck's attempt to relate concepts of different origins and from different levels, gives rise to complex assumptions. In order to determine the extent to which these assumptions could generate important findings, it is proposed to discuss the material pertinent to each deduction and experiment under the headings of theoretical assumptions, methodology and design, and statistical inferences.

(*a*) *Conditionability as a measure of reactive inhibition*

(i) *Theoretical assumptions*

In his paper 'A dynamic theory of anxiety and hysteria' (1955*a*), and in his recent book (1957*b*), Eysenck developed and set out two postulates which were derived by combining certain observations by Pavlov (1927) with Hull's formalized construct of reactive inhibition I_R (1943). The first postulate states that there are individual differences with respect to the speed with which reactive inhibition is produced and dissipated and with respect to its strength, the differences themselves being properties of the physical structures involved in the evocation of responses. The second, typological postulate states that strong reactive inhibition and its dissipation are linked to extravert behaviour patterns and hysteric/psychopathic disorders while weak inhibition is said to be associated with introversion and dysthymic disorders.

It seems that in these postulates untested assumptions, hypotheses, and molar and molecular concepts are inextricably mixed. *The assumptions* are that reactive inhibition is identical with the Pavlovian concept of internal inhibition, that the process is neurologically determined, and that there are specific neurological differences between anxious and hysteric patients. *The hypotheses* link the process of reactive inhibition with extraversion and hysteria, assuming the validity of the introversion-extraversion dimension. No attempt is made to differentiate between Hull's largely molar theory and the molecular basis of the Pavlovian concept. Implied in the nature of the experimental situations are a specific neurological theory and mechanism which Eysenck (1955*a*, 1957*b*) does not specify, and even rejects, yet precise deductions are made on the basis of it. An attempt will be made to show that the assumptions are unwarranted and untested, so that the hypotheses derived from them are reduced to the status of hunches.

Eysenck takes his original cue from Pavlov (1927) that:

> ...a direct transition from an inhibitory to an excitatory rate of stimulation of the same place of the skin...may lead to different forms of disturbance, depending on the type of nervous system of the animal. In dogs with the 'more resistant' nervous systems it leads to a predominance of excitation, in dogs with the 'less resistant' nervous system to a predominance of inhibition. So far as can be judged on the basis of casual observation I believe that these two variations in the pathological disturbance of the cortical activity in animals are comparable to the two forms of neurosis in man—in the pre-Freudian terminology neurasthenia and hysteria—the first with exaggeration of the excitatory and weakness of the inhibitory process, the second with a predominance of the inhibitory and weakness of the excitatory process.

This statement prompts Eysenck (1955*a*) to '...formalize...this acute observation... in terms of a general psychological law proposed by Hull namely his law of reactive inhibition'.

There is, of course, no evidence beyond their semantic affinity to suggest that the two uses of the word inhibition cover identical processes. This should be a separate hypothesis. Eysenck seems to assume their identity, however, when he continues (1955*a*, p. 34): 'It is easy to see what Pavlov and Hull had in mind in advocating this process of inhibition.' Yet Pavlov's theorizing involved a neurological model so that cortical inhibition is interpreted as an intervening variable, while Hull's concepts, though influenced by the Pavlovian model, are conceived on quite a different level as mainly hypothetical constructs. While it is conceivable that cortical and reactive inhibition may be different names for the same neurological processes, their identity cannot be tested simply by making this assumption. It is clear that Eysenck (1957*b*, p. 53) does not differentiate between intervening variables and hypothetical constructs. This may be one of the reasons why his postulates contain at least elements of an attempt at the unification of theories at the semantic level or in terms of analogies, a procedure for which he berates American investigators (cf. Eysenck, 1957*b*, pp. 85 and 105).

Eysenck's first prediction on the basis of his second postulate may be examined now in terms of the above considerations. It runs as follows: '(in)...individuals in whom the excitation–inhibition balance is tilted in the direction of strong excitation and weak inhibition, conditioned responses should be formed quickly and easily and should be difficult to extinguish. Conversely, in individuals in whom the excitation-inhibition balance is tilted in the direction of weak excitation and strong inhibition, conditioned responses should be formed slowly and with difficulty and should be easy to extinguish' (Eysenck, 1955*a*, 1957*b*). For this prediction to be borne out, there should be a valid measure of introversion-extraversion in terms of which excitation-inhibition balance is defined, and the hysteric group should be marked by greater spontaneous recovery, slower rate of C.R. acquisition, as well as by lower total frequency with which the C.R. is elicited. Over and above such requirements, it will be recognized that these predictions make a number of assumptions concerning conditioning theory which are not universally supported. Though Eysenck is at pains to avoid the limitations of any specific neurological model, the impression would seem to be conveyed that he accepts certain aspects of a Pavlovian model of conditioning by implication, and conditioning and extinction are accounted for in terms of this theory. A Pavlovian model of conditioning is, of course, not generally acceptable (Eccles, 1953; Gellhorn, 1953). Eysenck also appears to assume invariability in the establishment of C.R.'s of the autonomic nervous system irrespective of whether they are alimentary or defensive responses. The available evidence would suggest (Razran, 1956) that defensive C.R.'s may be established more quickly than alimentary ones, and furthermore, that under certain conditions a response that is easily acquired, may be easily extinguished.

(ii) *Methodology and design*

Eysenck (1955*a*, 1957*b*), considers that Franks's experiments (1954, 1956) have fully confirmed his first prediction from his second postulate. For the purpose of this part of the discussion it will now be assumed that the assumptions underlying Eysenck's postulates are justified and that the conditionability hypothesis represents their valid and logically correct development. To determine the validity of Eysenck's claim, it is necessary to look more closely at the method and design of Franks's (1954, 1956) conditioning experiment. This investigation was designed to test the hypothesis that dysthymics condition more quickly and more strongly and extinguish less quickly than either hysterics or normal controls.

Fundamental objections may be raised against the use of an unconditioned stimulus (puff of air to the eye) of constant intensity for all experimental groups (pressure of 65 mm. of mercury). If, as in the present case, neurological differences between the experimental groups are assumed, it follows that the same stimulus has a different sensory value for hysterics and dysthymics. Thus one would predict that dysthymics who may have low sensory thresholds would show stronger avoidance reaction to the air puff than hysterics who may have high sensory thresholds and therefore less marked avoidance reaction. In other words, differences found by Francks (1954, 1956) may be due to differences in sensory thresholds, and to differences in subjective stimulus value rather than to differences in conditionability. In order to confirm that true differences in conditionability did exist between Franks's experimental groups, absolute thresholds for the air puff should have been determined for each subject as was done in the case of the conditioned stimulus (buzz) for less weighty reasons. Assuming again that individual differences in reactive inhibition are identical with individual differences in excitation-inhibition balance, differences in reactive inhibition as described by differences in conditionability should have been determined by presenting unconditioned stimuli of an intensity that is a simple multiple of each individual's threshold in the relevant eyeball receptors.

The graphs presented by Eysenck (1955*a*, 1957*b*) and Franks (1956) indicate that at the beginning of the conditioning experiment there was already a large difference in strength of C.R. acquisition between dysthymics and hysterics. It is held here that the data from the experimental groups are comparable only if they start from a common base-line, or if comparable points of origin are first obtained. One question that Franks's design unfortunately leaves unanswered is whether his hysteric group would have conditioned better if the intensity of the unconditioned stimulus had been increased. Franks himself writes: 'It was noticed that an air puff at a pressure of only 65 mm of mercury was in several cases insufficient to produce even an unconditioned P.G.R. after the first few trials' (1956, p. 147), though the subjects who were thus resistant were not specified. The conditioning study furthermore did not include the vital analysis of differences in the rate of conditioning, i.e. differences in the slopes of the curves. A visual inspection of the curves strongly suggests that the difference between hysterics and dysthymics is not significant. (Data presented by Eysenck (1957*b*) based on the fitting of exponential curves to the results of Franks's (1956) conditioning experiment would seem to leave the position ambiguous.) In order to obtain all the necessary evidence for the reactive inhibition hypothesis, Franks's investigation should also have included testing for spontaneous recovery.

The method of selection of subjects in Franks's study raises certain difficulties because of selection by a double criterion. In addition to the clinical diagnosis of hysteria or dysthymia, Franks required a score of 33 or greater on Guilford's Rhathymia scale for experimental hysterics, and a score of 32 or less for dysthymics. This procedure has several implications. It could mean, for example, that psychiatric diagnosis is valid for hysterics scoring above 33 on the scale and for dysthymics scoring below 32, but that it is invalid for the omitted subjects. Basically, the procedure assumes that the R scale is a valid test of normal and neurotic extraversion. Even assuming its validity, it still remains true that selection by the R scale may be only in terms of *social* introversion or extraversion with a large error of mis-classification and the disadvantage of confining the experimental groups to very homogeneous and unrepresentative samples of the neurotic subgroups. While the selected groups no doubt contained an unspecifiable number of genuine hysterics and dysthymics, it is clear that any differences in conditionability between the two groups are primarily related to scores on the R scale and whatever that scale may be said to measure. The general relationship between introversion-extraversion and conditionability can only remain equivocal, while the relationship between excitation-inhibition, reactive inhibition and anxiety and hysteria remains to be tested.

Finally, the hypothesis required the exclusion of the alternative hypothesis that conditionability of the eyeblink response is related to degree of neurosis or neuroticism. This required experimental groups matched for severity of illness which clinically is a nearly impossible task, while in terms of objective tests no unidimensional measures of neuroticism were available to Franks. No conclusions with regard to the relationship between conditionability and degree of neurosis or neuroticism can therefore be reached. A further drawback of Franks's study is that his experimental groups were not equated for age, intelligence or sex. This omission leaves the door wide open to alternative explanations of the findings.

(iii) *Statistical inferences*

On the basis of Franks's findings Eysenck (1955*a*, 1957*b*) concludes that his own predictions were borne out by the facts. There is insufficient evidence to sustain this conclusion for the following reasons. (1) The assumptions underlying Eysenck's postulates were not directly tested by his hypotheses. (2) No unambiguous differences between rates of C.R. acquisition and extinction of hysterics and dysthymics were demonstrated, and there was no evidence of spontaneous recovery. (3) Although there were significant correlations between the R scale and conditionability of the eyeblink response, the reliability of acquisition and extinction trials calculated for the normal group were sufficiently low (0·52 and 0·37 respectively) to cast serious doubt on the repeatability of the findings. (4) The correlation between the R scale and conditioning as measured by eyeblink and P.G.R. (a notoriously unreliable variable), is sufficiently low to leave a substantial error variance to be accounted for. (5) Eysenck presents no evidence that the two neurotic groups were equally neurotic so that the question as to the relatedness of neuroticism and conditionability is left unanswered. (6) The composition of the experimental groups in terms of age, sex and intelligence is significantly different so that alternative hypotheses are appropriate.

Table 3. *Data of Franks's groups*

	Normals	Hysterics	Dysthymics
Mean age	24·3	28·8	31·3
Mean scores Progressive Matrices (1938)	57	44·6	50·2
No. of female subjects	14/20	13/20	5/20

Table 3 summarizes the age, sex and intelligence data in Franks's (1954) study. Normals are significantly younger than either hysterics or dysthymics. Chi square for differences in sex between the groups is significant at the 2% level, and all '*t*' tests between the groups for differences in intelligence are significant. Correlations between intelligence and conditionability are said by Franks (1956) not to be significant. It would appear that he used in this case a two-tail test of significance, on which the correlations between conditionability and intelligence are consistently positive, frequently approaching significance. It would seem that two-tail tests of significance are used almost arbitrarily depending on the mere technicality of hypothesis formation. Scientific method requires that formulation of an hypothesis should be a direct function of the strength of the theoretical framework on which it is based and the background of data supporting it. There would appear to be as much (or as little?) justification for linking by hypothesis conditionability and introversion, hysteria and Rhathymia, as for linking intelligence and conditionability. One might further be tempted to argue that in the field of personality study only two-tail tests of significance provide an adequate criterion of confidence. An alternative hypothesis in terms of intelligence may well account for Franks's differences between hysterics and dysthymics. Selective sampling would lend support to this conclusion. For example, more than half the groups of dysthymics were obsessionals who are generally found to be more intelligent than other neurotics. Also, Franks excluded all subjects with an I.Q. of less than approximately 90 on the Wechsler–Bellevue Scale. Since hysterics are considered to be on the average of lower intelligence than anxious or obsessional patients, sampling may have played an important part in Franks's findings. Evidence inconsistent with an intelligence hypothesis is found in the wide differences in Matrices I.Q.'s between hysterics and normals, and in the general absence of correlations between autonomic functions and intelligence. It may, therefore, be necessary to frame an alternative hypothesis in terms of intelligence, as well as of age and sex, to account for some of the similarities in conditioning between normals and hysterics.

A more recent study of Franks (1957) would appear to substantiate his and Eysenck's conclusions on the basis of the earlier conditioning study. In this later study a significant negative correlation between conditionability of the eyeblink response and extraversion in a non-neurotic group was found. The measure of extraversion in this case was the Maudsley Personality Inventory in which the extraversion (E) scale was developed from Guilford's inadequate Rhathymia scale. In a validation study with neurotic introverts and extraverts it has now been established (Franks, personal communication, 1957) that the E scale of the M.P.I. is not a valid measure of extraversion.

The conclusion would seem to be inevitable that Eysenck does not have the necessary experimental data to support his postulate concerning differences in excitation-inhibition, or differences in reactive inhibition, as demonstrated in differences in conditionability of the eyeblink response, relatable to differences in introversion-extraversion or anxiety and hysteria.

(b) *Kinaesthetic figural after-effect as a measure of reactive inhibition*

(i) *Theoretical assumptions*

Eysenck's second prediction from his second postulate is in terms of figural after-effect phenomena. It states: ‘...that figural after-effects should appear earlier among extraverts, they should appear more strongly in extraverts and they should disappear more slowly...as compared with introverts’ (1955*a* p. 38). Also implied by the theoretical framework are predictions that hysterics should have more pronounced reminiscence effects than dysthymics and that in experiments with reversible perspectives hysterics should show a greater number of reversals per time unit as well as a greater progressive rate of reversals compared with dysthymics.

Eysenck (1955*a*) deals at some length with figural after-effect phenomena, so that this particular deduction will have to be considered in some detail. In formulating his hypothesis concerning differences in kinaesthetic figural after-effect co-varying with introversion-extraversion, he makes the following assumptions (these he justifies elsewhere in some detail (1957*b*) and his modified approach will be considered below):

(*a*) ‘...“satiation” (Köhler & Wallach, 1944) appears to correspond exactly to the Pavlov–Hull notion of “inhibition”’ (Eysenck, 1955*a*, p. 38). (*b*) The processes determining visual figural after-effect and the experimental procedure eliciting it correspond to the processes and procedures in an experiment on kinaesthetic figural after-effect. (*c*) Neuro-physiological considerations and details are unnecessary and not ‘...strictly relevant to a psychological theory of the kind developed here, which remains throughout at the molar level and is not concerned with physiological details’ (Eysenck, 1955*a*, p. 38).

The last assumption indicates that Eysenck wishes to dissociate himself from Köhler's (Köhler & Wallach, 1944) electrotonic theory, in fact from all neuro-physiological models, and that he wants his own theory to remain at the molar level. Despite that he keeps close to the analogy of Pavlovian excitation/inhibition and would thus appear to retain specific hypotheses concerning neural mediation. It would seem to be all too easy to equate cortical and reactive inhibition and satiation simply because some form of inhibition appears to be involved in each. In general it would seem that a theory is being presented that requires a reasonably precise definition of its parameters which would allow the equating of inhibition and satiation, but none is offered.

It would also seem that Eysenck leans all too heavily on the experiments by Klein & Krech (1952). These authors, using the test of kinaesthetic figural after-effect, described in detail by Eysenck (1955*b*, 1957*b*), with brain-damaged patients, claim that their data '...support the hypothesis that one effect of cortical lesions is marked alterations in general cortical conductivity'. It is, of course, evident that this hypothesis was not directly tested. All that can be said for the conclusion is that it is correct provided that the original assumption was correct. Klein & Krech would seem to realize the tenuousness of their conclusion when they write 'Future work should explore the possibility of direct physiological measurements of conductivity to replace the indirect psychological ones used in this study'. In a later article, however, at least one of the authors (Livson & Krech, 1955) is in favour of testing further for behavioural concomitants of what is termed cortical conductivity. The discussion in this paper on what is recognized to be a very speculative theory shows, however, much greater awareness, by contrast to Eysenck, of the tenuous nature of cortical constructs and theoretical frameworks built around them. It would, furthermore, now appear that any hypothesis linking cortical conductivity and satiation phenomena must take into account the study of Brenner (1956), who found no differences between brain-damaged and normal groups in their perception of apparent movement after continuous stimulation. According to Eysenck's theory (1955*a*, 1957*b*) such differences should have been found, because he considered that brain damage leads to an increase in extraversion, because of increases in cortical and reactive inhibition.

So far, therefore, the identity of reactive inhibition, conceived as a variant of cortical conductivity by Eysenck, and satiation as defined by Köhler, remains an assumption for which there is little positive evidence. This assumption has been incorporated into the hypothesis dealing with the relationship between satiation or figural after-effect and extraversion which may thus be said to have lost a great deal of its efficiency and validity. It may be thought that a stronger case might have been made by considering the affinity of Pavlov's concepts of 'less or more resistant nervous systems' and the figural after-effect phenomena. I_R in Hull's system is primarily an explanatory concept in drive reduction/reinforcement, and its specific neurological implications in Eysenck's theory may deprive a valuable concept of a great deal of its explanatory powers.

In his latest book (1957*b*), Eysenck refers (pp. 151–54) to three possible proofs of the identity aspect of the postulate linking reactive inhibition, satiation and typological differences: neuro-physiological, functional and direct and indirect differential proofs. Differential proofs are usually conveyed statistically by correlations between co-varying parameters, the difference between the direct and the indirect versions depending on whether the relationship between two parameters is measured by correlating several tests representing the postulated relationship say between reactive inhibition and satiation, or by postulating a third parameter to which both are said to be related. In concrete terms, if satiability is correlated with extraversion, and reactive inhibition as represented by conditioning or rather the lack of it is also related to extraversion, then this may be taken as 'proof' of the identity of reactive inhibition and satiation processes. This is clearly a fallacious argument. A few simple Euler type of diagrams would show that the indirect differential proof is acceptable only if the parameters or functions A and B described by the correlation between say A and C and between B and C are identical. It is furthermore dependent on the size of the correlations obtained. In view of the ambiguous relationship between conditionability, figural after-effect and introversion-extraversion this seems an extremely risky inferrential procedure which it is difficult to justify either in scientific method or on economical grounds. Since Eysenck himself admits (1957 *b*) that the tests used must be reliable and valid and that the obvious irrelevant variables of age, sex and intelligence must be controlled, it must be concluded that in the present theoretical context this type of proof is not acceptable because of the deficiencies on just these counts of the various studies that have been considered.

Finally, what of the assumption that the processes of visual and kinaesthetic figural after-effect are identical? Köhler's findings are that figural after-effect (visual) depends on intensity of stimulation as much as on duration over time. In experiments with kinaesthetic figural after-effect duration over time has been the independent variable. Köhler's theory furthermore is quite specific concerning the asymptotic character of satiation curves (1944, p. 330), that displacement of a test object depends upon its position with regard to the affected area, i.e. the area of maximum satiation (1944, p. 337), and that satiation is cumulative and may depend on sensory experience prior to experimentation (1944, p. 341). In the studies on kinaesthetic figural after-effect that are considered here (Eysenck, 1955 *b*; Nicholls, 1956), intensity of stimulation and prior sensory stimulation were not controlled. (Parenthetically, it may be considered difficult to decide whether figural after-effect or satiation as defined by Köhler were in fact measured in a bilateral experiment.)

(ii) *Method and design*

Sufficient has been said to indicate that the assumptions behind Eysenck's hypothesis concerning satiation do not appear to be supported by sufficient empirical evidence. The construction of the two experiments considered here (Eysenck, 1955 *b*; Nicholls, 1956) raises three questions, two of which have already been discussed.

The first of these concerns the selection of subjects. A double criterion was employed as in Franks's (1956) study though different cut-off points on the R scale were used by Eysenck (1955 *b*). As before, any correlation between this scale and the figural after-effect would make it difficult to consider as established a relationship between anxiety and hysteria and satiation. In Nicholls's study this difficulty was overcome by using clinically diagnosed groups of hysterics and dysthymics. The second difficulty in the design arises from the lack of control of pre-experimental stimulation. If Köhler's theory is correct, one would expect that the socially extraverted hysterics might have commenced the experiment partially satiated (Hamilton, 1956 *b*). Eysenck (1957 *b*) seems to have accepted the possibility that pre-experimental satiation of hysterics may have affected his own and Nicholls's results. His subsequent suggestion that a lengthy period of dark adaptation might reduce this semi-permanent satiation to vanishing point would not seem a very apt one in relation to the kinaesthetic figural after-effect.

The third question of methodology concerns the intensity with which subjects rubbed the stimulus object prior to judgement. This according to Köhler (Köhler & Wallach, 1944) is one of the two crucial parameters. In both the studies considered here it was left uncontrolled. It may be held that the less timid hysteric, selected because of his high score on social extraversion, would employ a firmer finger grip on the apparatus. In terms of Köhler's theory he would thus be expected to show a more pronounced figural after-effect than the diffident dysthymic.

(iii) *Statistical inferences*

Considering his satiation experiments, Eysenck (1955*a*) claims that '. . . our prediction is borne out at a statistically significant level in every detail'. This conclusion derives only little support from Eysenck's own study (1955*b*), and may be considered to be invalidated by Nicholls's (1956) study. The main reasons for considering Eysenck's own study as inconclusive are as follows. (1) An overall test of significance of difference between hysterics and dysthymics was significant, but only one '*t*' test in nine reached significance when considering the separate stimulation periods. (2) Only two in six correlations between the R scale and kinaesthetic figural after-effect were significant. (3) Intensity of stimulation and pre-experimental stimulation were not controlled, and the percentage decrement curve (Eysenck, 1955*a*, *b*) approached the asymptote in only one case, i.e. some major requirements of Köhler's theory were not met. On the basis of Eysenck's study, therefore, it would not be safe to say more than that there appears to be a relationship between kinaesthetic figural after-effect and high scores on the R scale, which would need to be more analytically investigated. No relationship between anxiety and hysteria, reactive inhibition, satiation and introversion-extraversion has been conclusively demonstrated by this study.

Nicholls (1956) set out to test the generality of figural after-effect in twenty different stimulus situations, employing clinically diagnosed hysterics and dysthymics as subjects. He failed to find any differences in figural after-effect between his groups except in one out of three settings of the kinaesthetic figural after-effect apparatus, and in a figure reconstruction test which did not correlate with kinaesthetic figural after-effect. No, significant correlations between the R scale and the measures of figural after-effect were found, and the data contra-indicated the generality of the variable or construct measured by figural after-effect and similar tests. In this study, hysteria and the R scale were not linked by hypothesis so that a two-tail test of significance was appropriate. This fails to reach statistical significance, though '*t*' (1·762, D.F. = 50) is significant at better than the 5% level on a one-tail test. With these findings Eysenck's hypothesis linking hysteria, reactive inhibition and cortical satiation finds little support.

Eysenck (1957*b*) attempts to account for the negative findings of Nicholls's experiment by suggesting that the hysterics' inability to fixate for sufficiently long in experiments of visual figural after-effect led to a reduction in the build-up of reactive inhibition or satiation. This would seem to be a qualitative observation and Eysenck does not present empirical data to support his alternative hypothesis. If the postulated distractibility had been empirically demonstrated, it may well be thought surprising that it was not considered to be a disturbing factor in the kinaesthetic experiments. Difficulties over the hysteric sub-group in these experiments need not have invalidated all the data, because the postulated relationship in terms of a negative correlation between introversion and figural after-effect might have been established. Eysenck does not indicate whether this analysis had been made.

With respect to other predictions implied by Eysenck's (1955*a*) postulates, similar negative evidence is forthcoming. The relationship postulated between reminiscence effects and introversion-extraversion (Eysenck, 1955*a*) has now been shown to be only very limited and slight in terms of correlations for normal extraverts (Eysenck, 1957*b*, p. 126), while a relationship between neurotic introversion-extraversion and reminiscence has not been found (Eysenck, 1957*b*, footnote, p. 254). As in the case of the figural after-effect experiments, Eysenck holds the distractibility of hysterics responsible, resulting from a need to dissipate reactive inhibition. He does not consider it necessary to re-examine his deductions from his basic postulates or the justification for the postulates. Turning to predictions about oscillations of reversible perspectives, a study by Hamilton (1957*b*), substantiated by Nicholls (1956), shows that hysterics have the smallest number of reversals, and no marked differences between hysterics and other neurotics were found in changes in rates of oscillation. (Such changes would, for example, appear in lowered reliability coefficients for reversible perspectives calculated by the split-half method applied to the total inspection period.)

IV. CONCLUSIONS

An analysis of Hildebrand's study indicated that the major measures of extraversion, in terms of their factor loadings on the introversion-extraversion factor, were deficient in unidimensionality and in validity, and that the R scale particularly was an unreliable measure of pathological extraversion, i.e. hysteria.

In view of the marked extraversion—in terms of the tests—of Hildebrand's normal control group, it would be difficult to agree with Eysenck that he is justified in using dysthymics and hysterics as criterion groups for the introversion-extraversion dimension, or that the existence of this dimension has been formally proved. When Eysenck writes that the dimension can be operationally defined in terms of a few objective tests, he would appear to reduce artificially the complexity of a personality type. It is almost platitudinous to say that an hysteric is characterized by perhaps much more important forms of behaviour or mechanisms than those expressed by scores on Guilford's Rhathymia scale or the Porteus Mazes. Eysenck's own conception of the structural hierarchy of personality (1947) strongly suggests that he, too, does not really believe that a type is adequately defined in terms of a few data at the habit or trait level.

The second part of the review dealt with hypotheses deduced from postulates linking introversion-extraversion, anxiety and hysteria with differences in reactive inhibition. The basic assumptions underlying the postulates were critically examined. They were held to rest on insufficient evidence. The methodology applied to the testing of the hypotheses linking anxiety and hysteria, introversion and extraversion with conditionability and cortical satiation, was considered to be inadequate to sustain Eysenck's conclusions. It is, of course, likely that most rapid progress in the development of genetic theories of personality may be made by considering hypotheses derived from different levels of bio-psychological functioning. In view of the degree of uncertainty, however, that *ipso* *facto* attaches to any assumptions supporting inter-level hypotheses it would probably pay long-term dividends either to limit the assumptions, or to submit them to prior testing. Much wasted labour may be avoided if hypotheses could be derived from postulates in which the assumptions have maximal empirical support. This would appear to be all the more necessary when one compares the relative grossness of Eysenck's typological postulate with the detailed and specific predictions derived from it. It is apparent that Eysenck (1957*b*) has realized the need to limit the extent of precise quantitative predictions in relation to the principles involved in his postulates. Whereas previously (Eysenck, 1955*a*) the stress has been on the preciseness of the predictions possible in his postulate system, Eysenck now (1957*b*) stresses the qualitative aspect of the postulates '...derived in a non-rigorous way from previous experiments and generalizations...'. In this 'mongrel type of theory' (1957*b*, p. 251) predictions are frequently '...much less rigorous and verification is either partially or quite non-existent'. It would appear to be open to discussion now whether Eysenck's new theoretical approach remains at the hypothetico-deductive level. Furthermore, lack of verification of deductions can be laid more firmly at the door-step of 'non-vigorously' derived postulates, rather than inadequate experimental techniques, sampling, or other *ad hoc* explanations of negative findings.

The only possible justifiable conclusion on the findings of the conditioning and figural after-effect experiments discussed here is that some relationship exists between these two mechanisms and the behavioural qualities measured by the R scale. It is likely, considering the data of Table 1, that these qualities may be subsumed under the concept of social extraversion. If this were independently substantiated, the absence of discriminatory ability of this scale may become explicable.

It is thus difficult to share Eysenck's optimism concerning the experiments that have been discussed, an optimism that finds its expression in the elucidation of a 'Dynamic theory of anxiety and hysteria' (1955*a*). Eysenck writes there (p. 40) that the experiments on conditionability and kinaesthetic figural after-effect '...were carried out especially in order to provide evidence regarding the power of our theory to make possible testable deductions'. It is, of course, valuable to have a theory formulated in testable terms and with such specificity that it becomes amenable to experimental verification. It is generally agreed that psychiatric and psychoanalytic theories as stated at present are too comprehensive and have too many provisos to be directly testable. That makes the theories experimentally untested ones, but not demonstrably wrong. Eysenck's 'Dynamic theory of anxiety and hysteria' is an ingenious attempt to account for an enormous variety of psychological data parsimoniously—perhaps too parsimoniously. The deductions from the theory in relation to reactive inhibition as the 'dynamic variable' do not appear to have been verified to any considerable extent, so that in terms of this most important criterion it is not a good theory.

This paper is based on a dissertation submitted in part fulfilment of the Academic Postgraduate Diploma in Abnormal Psychology of the University of London, 1956, while the writer was a postgraduate student of the University at the Institute of Psychiatry 1955–56. He is indebted to Mr Harry Hurwitz, Birkbeck College, and Mr Gwynne Jones, Institute of Psychiatry, of the University of London, who critically read parts of the manuscript.

References

Albino, R. C. (1953). Some criticism of the application of factor analysis to the study of personality. *Brit. J. Psychol.* **44**, 164–8.

Brenner, M. W. (1956). The effects of brain damage on the perception of apparent movement. *J. Personality*, **25**, 202–11.

Christie, R. (1956*a*). Eysenck's treatment of the personality of Communists. *Psychol. Bull.* **53**, 411–30.

Christie, R. (1956*b*). Some abuses of psychology. *Psychol. Bull.* **53**, 439–51.

Eccles, J. C. (1953). *Neurophysiological basis of the mind.* Oxford: Clarendon Press.

Eysenck, H. J. (1947). *Dimensions of Personality.* London: Kegan Paul.

Eysenck, H. J. (1953*a*). *The Structure of Human Personality.* London: Methuen.

Eysenck, H. J. (1953*b*). The application of factor analysis to the study of personality: a reply. *Brit. J. Psychol.* **44**, 169–72.

Eysenck, H. J. (1955*a*). A dynamic theory of anxiety and hysteria. *J. Ment. Sci.* **101**, 28–51.

Eysenck, H. J. (1955*b*). Cortical inhibition, figural after-effect and theory of personality. *J. Abnorm. Soc. Psychol.* **51**, 94–106.

Eysenck, H. J. (1957*a*). Drugs and personality. I. Theory and methodology. *J. Ment. Sci.* **103**, 119–31.

Eysenck, H. J. (1957*b*). *The Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

Franks, C. (1954). An experimental study of conditioning as related to mental abnormality. Unpublished Doctor's dissertation. University of London Library.

Franks, C. (1956). Conditioning and personality: A study of normal and neurotic subjects. *J. Abnorm. Soc. Psychol.* **52**, 143–50.

Franks, C. (1957). Personality factors and the rate of conditioning. *Brit. J. Psychol.* **48**, 119–26.

Gellhorn, E. (1953). *Physiological Foundations of Neurology and Psychiatry.* Minneapolis: University of Minnesota Press.

George, E. I. (1954). Experimental study of the relation between attitudes and personality. Unpublished Doctor's dissertation. University of London Library.

Hamilton, V. J. (1956*a*). Reaction to uncertainty and ambiguity as a personality variable. Unpublished Doctor's dissertation. University of London Library.

Hamilton, V. J. (1956*b*). Eysenck's dynamic theory of anxiety and hysteria—a methodological critique. Unpublished dissertation for the Diploma in Abnormal Psychology, University of London.

Hamilton, V. J. (1957*a*). Conflict avoidance in obsessionals and hysterics and the validity of the concept of dysthymia. *J. Ment. Sci.* **103**, 666–76.

Hamilton, V. J. (1957*b*). Perceptual and personality dynamics in responses to ambiguity. *Brit. J. Psychol.* **48**, 200–15.

Hanley, Ch. & Rokeach, M. (1956). Care and carelessness in psychology. *Psychol. Bull.* **53**, 183–6.

Hildebrand, H. P. (1953). A factorial study of introversion-extraversion by means of objective tests. Unpublished Doctor's dissertation. University of London Library.

Hull, C. L. (1943). *Principles of Behaviour.* New York: Appleton-Century.

Klein, G. S. & Krech, D. (1952). Cortical conductivity in the brain-injured. *J. Personality,* **21**, 118–48.

Köhler, W. & Wallach, H. (1944). Figural after-effects: An investigation of visual processes. *Proc. Amer. Phil. Soc.* **88**, 269–357.

Livson, N. & Krech, D. (1955). Dynamic systems, rote learning, and retroactive inhibition. *J. Personality*, **24**, 2–19.

Martin, I. D. (1955). An experimental study of muscular tension in psychiatric patients. Unpublished Doctor's dissertation, University of London Library.

Nicholls, G. (1956). The relationship between certain personality variables and figural after-effect Unpublished Doctor's dissertation. University of London Library.

Pavlov, I. P. (1927). *Conditioned Reflexes.* London: Oxford University Press.

Razran, G. (1956). Extinction re-examined and re-analysed: A new theory. *Psychol. Rev.* **63**, 39–52.

Rokeach, M. & Hanley, Ch. (1956). Eysenck's tender-mindedness dimension. *Psychol. Bull.* **53**, 168–76.

(*Manuscript received* 13 *September* 1957)

ANXIETY AND HYSTERIA—A REPLY TO VERNON HAMILTON

By H. J. EYSENCK

Institute of Psychiatry, University of London

In the introduction to my book, *The Dynamics of Anxiety and Hysteria*, I pointed out that: 'To indicate the tentative nature of the hypotheses put forward, I have included in the last chapter a discussion of various criticisms to which the general theory here put forward is subject; this should have the added advantage of enabling reviewers to pass straight from the introduction to the conclusion without having to read the intervening parts.' Hamilton, in his 'methodological critique', has apparently followed this advice, which was not meant to be taken seriously; he has also taken up many of those criticisms which I have put forward myself, and which he now presents as his own. In addition, he has quite misinterpreted a number of facts in a way which is difficult to reconcile with the care required of anyone purporting to examine any body of scientific work. In dealing with his points I shall concentrate on the main issues, leaving aside the many statements of opinion which fill the pages of Hamilton's essay.

To begin with, Hamilton contends that 'Eysenck's theory of the relationship between the dimension of introversion-extraversion and the current nosological categories of the neuroses is primarily based, not on his own earliest investigations, but on the findings of Hildebrand'. He then goes on to criticize Hildebrand's Ph.D. thesis on a number of points. It is interesting to note the number of errors in this one sentence. The theory in question, as I have always been careful to point out, is due to Janet and Jung, not to myself. What I have done has been simply to make deductions from their hypotheses and to test these experimentally. These deductions are of several different kinds, ranging from factor analytic to more strictly experimental ones. Hildebrand's work is only one of many strands; if it had never been carried out, the evidence available would still have been quite sufficient, in my view, to support Jung's hypothesis. The exclusive concentration on Hildebrand's study is therefore quite misplaced.

The actual criticisms made by Hamilton of the Hildebrand study illustrate a lack of awareness of even the more simple theorems of factor analysis. Thus, in discussing his Table 1, he claims 'that the major measures of introversion-extraversion and of neuroticism are significantly correlated. These results make it clear that the Guilford scales cannot produce an undiluted measure of either neuroticism or of introversion extraversion and that orthogonal solutions of such a matrix are factor analytical artifacts rather than evidence for the unidimensionality of Eysenck's measures.' This is quite erroneous. It occurs quite frequently that factors are defined by tests which also have some variance on another factor. This merely indicates that the particular test in question is not univocal; it does not mean that the factors themselves are not orthogonal. Indeed, the term 'factor analytical artifact' in Hamilton's argument is meaningless as it stands. All statistical and methodological analyses are 'artifacts', i.e. 'a product of human art and workmanship': it is difficult to see how they could be anything else.

Another statistical curiosity is Hamilton's argument that 'the so-called abnormally high score of Hildebrand's original control group is probably no more than 1 S.D. above the mean for tested normal control groups'. Anyone finding that a group to be used as an average control group for the purpose of intelligence testing had a mean I.Q. in excess of 116 would presumably conclude that something odd had happened to his sampling, and would be justified in substituting a group with a mean nearer 100. Why this should not be permissible here, where the derivation of the group itself strongly suggested a somewhat psychopathic-extraverted character, is not quite clear to me. It might be argued that Hildebrand ought to have shown more circumspection in finding his control group; that seems to be a reasonable suggestion. It does not seem reasonable to suggest that a mistake, once made and clearly recognized, should never be made good, particularly when, as in this instance, objective evidence is available about the non-representative nature of the group concerned. This can hardly be called an *ad hoc* argument in view of the fact that a highly significant difference is found in R scores between the sample presumed to be normal, and the known British population values. Having admitted on one page that Hildebrand's normal control subjects were about 1 S.D. more extraverted (on the R scale) than large random samples in which the test was standardized, Hamilton can hardly be serious in saying on the next page that 'a comparison of Hildebrand's normal controls with other control subjects tested on the same scale does not support Eysenck's and Hildebrand's special plea that Hildebrand's normal control subjects were abnormally extraverted'. A 'special plea' can seldom have had such excellent support!

Hamilton goes on to attack the use of the Guilford R scale, and of the Extraversion Scale of the Maudsley Personality Inventory, based in part on the R scale, as measures of extraversion. His argument confuses two quite distinct points, and as it is very important that they should be kept clear, I shall discuss this point in detail. Jung's hypothesis states that 'much the most frequent neurotic disorder of the extraverted type is hysteria...'. On the other hand, speaking of the introvert, he maintains that 'his typical neurotic disorder is psychasthenia', or as we would nowadays say, dysthymia. Two deductions may be made from these statements. The first one is that *on any good measure of extraversion, dysthymics should have lower scores than hysterics or psychopaths.* (The term 'hysterics', as used by psychiatrists at the time when Jung wrote his famous book, included people whom we would now call psychopaths; this will be clear from the quotations I have given in chapter 6 of my book.) This is absolutely essential, and indeed has been found to be true by Hildebrand of the R scale. It has also been found to be true of the M.P.I. Extraversion Scale in a study by Sigal, Star and Franks, which Hamilton curiously enough quotes as disproving the value of this scale.* Others, such as Jensen and Claridge, have obtained similar results, and when all the available data are summed it appears that dysthymics have a mean score of 16, hysterics one of 25 and psychopaths one of 30, on the extraversion scale. The differences between the dysthymics and the other two groups are fully significant at the 1% level. These results provide strong support for Jung's hypothesis and for the adequacy of the scale in question.

The position is slightly more complicated when we consider the position of normal groups in relation to the neurotic ones. My own interpretation of Jung's hypothesis had always been that dysthymics would be more introverted, hysterics and psychopaths more extraverted, than a randomly selected normal group. The data available now confirm this prediction with respect to dysthymics and psychopaths, but not with respect to hysterics, who are only very slightly more extraverted than normals, whose mean score is approximately 24. This failure of the hypothesis may be linked with a curious feature relating to the intercorrelation of neuroticism and extraversion in our various groups. In normal samples this correlation tends to be around $-0{\cdot}1$, whereas in neurotic groups it rises to $-0{\cdot}4$. This cannot be accounted for in terms of selection. I thought at first that possibly the most neurotic and most extraverted groups might not be found in mental hospitals at all, but perhaps in prisons. However, studies with recidivist and other prison populations have shown that these have scores of both neuroticism and extraversion very similar to the scores obtained by hysterics. Furthermore, when we compared the correlations obtained from subgroups of our normal sample, selected for high and low neuroticism respectively, we found the same phenomena, i.e. zero correlations for the group which is low on neuroticism and substantial negative correlations for the group high on neuroticism. These relations are definitely extra-chance; they have been found quite independently by American investigators using the M.P.I. The explanation may be in terms of 'response set', or in terms of a multiplicative effect of emotionality and conditioning.

---

* Hamilton bases his view on a 'personal communication', not on published evidence. The actual scores obtained in this study were 21, 24 and 30 for dysthymics, hysterics and psychopaths. He does not explain how results so clearly in line with prediction can 'establish that the E scale of the M.P.I. is not a valid measure of extraversion'.

This problem is being actively pursued at the moment; fortunately its solution is of no particular importance in relation to Hamilton's criticism because the use of the R scale or the M.P.I. as a measure of extraversion is not dependent on the precise position of the normal group with respect to various clinical groups, but rather on the relative positions of the different clinical groups. Hamilton, in his discussion, mixes up these two points and is thus enabled to reach a conclusion which is not in accordance with the facts.

One further point should be mentioned. He complains that 'the overlap of all the data over the groups of controls, hysterics and dysthymics, is considerable'. I find it difficult to see what else one could have expected. In *The Scientific Study of Personality*, I have given a brief review of experiments into the reliability of psychiatric diagnoses, showing it to be rather low. When criterion reliabilities are well below 0·5, it is impossible even for a perfect instrument to give high correlations with the criterion, or to avoid 'considerable overlap' between the groups. It is because of this unreliability of psychiatric diagnosis that I have preferred, in working with neurotics, to use a double selection procedure by combining diagnosis and questionnaire score. Hamilton's criticism of this procedure is based on his belief that the questionnaire does not correlate with clinical diagnosis in the predicted manner. As this belief has been shown to be false, this criticism also falls to the ground.

We must now turn to Hamilton's discussion of my 'theoretical assumptions'. Having mis-stated my typological postulate (by leaving out all mention of excitatory potentials, which form an important part of it), he goes on to say that 'it seems that in these postulates untested assumptions, hypotheses, and molar and molecular concepts are inextricably mixed'. Later on he claims that 'the basic assumptions underlying the postulates were critically examined'. All one can say in answer to this is that Hamilton quite misconceives the function of postulates in scientific theory. A scientist is free to advance any postulate he wishes without having to give any reasons, without having to justify his assumptions or state his grounds for holding the postulate. There are only two restrictions on postulate making. In the first place the postulate must be specific enough to give rise to testable deductions or theorems, and in the second place these deductions, when tested, must not give rise to results incompatible with the postulate. As Hull points out in his *Principles of Behaviour*, 'scientific theory reaches belief in its postulates to a considerable extent through direct or observational evidence of the soundness of its theorems'. The whole discussion of scientific method given by Hull is apposite to Hamilton's critique because time and time again Hamilton complains that a particular test of my theory does not test the postulate *directly*. This, of course, is quite true; as Hull points out, such direct testing is impossible. All that we can do is test deductions and show that the 'empirical verification of theorems indirectly substantiates postulates'. Hamilton's failure to bear in mind these elementary principles of scientific methodology makes his discussion less valuable than it might otherwise be.

In turning from the postulates to the theorems and their verification, Hamilton rather arbitrarily limits his discussion to a few, while neglecting a wide range of other studies. This is an important point because the support which a postulate receives from experimentation is a direct function of the number of theorems tested. Furthermore, in dealing with those few theorems which he does discuss, Hamilton again arbitrarily limits this discussion to a few experiments. This, again, is not a useful method of procedure, because one's faith in the verification of a theorem rests in part on the number of studies successfully testing this particular deduction.

However, worst of all is Hamilton's tendency to misquote the research papers which he cites, and to give quite erroneous impressions on matters of fact. Thus he maintains that 'the graphs presented by Eysenck and Franks indicate that at the beginning of the conditioning experiment there was already a large difference in strength of C.R. acquisition between dysthymics and hysterics. It is held here that the data from the experimental groups are comparable only if they start from a common base-line, or if comparable points of origin are first obtained.' Even the most cursory reading of Franks's procedure would have shown Hamilton that exceptional care was in fact taken to arrive at a common base-line for the different groups, and that this endeavour was entirely successful. Hamilton's error arises from the fact that in the plotted curves the first point is a measure of the amount of conditioning on the first *test trial*. This test trial was preceded by several *conditioning trials*, and the superiority of the dysthymic group on the first test trial was precisely in line with prediction. Such carelessness in a 'methodological critique' is unusual; it becomes inexcusable when it is realized that this point was clarified by a letter sent in reply to an earlier draft of the paper.

Equally reprehensible is Hamilton's habit of making statements regarding the lack of availability of certain data when these data are in fact available. Thus, in connexion with Franks's conditioning experiment, he states that 'the hypothesis required the exclusion of the alternative hypothesis that conditionability of the eyeblink response is related to degree of neurosis or neuroticism'. He goes on to say that 'no unidimensional measures of neuroticism were available to Franks. No conclusions with regard to the relationship between conditionability and degree of neurosis or neuroticism can therefore be reached.' That this statement is simply not true, can be checked by anyone who will look at fig. 2, on p. 124, of volume 48, of *The British Journal of Psychology*, or at Franks's statement of results on p. 147, of volume 52, of *The Journal of Abnormal and Social*

Psychology: 'When the normals are compared with the neurotics as a combined group, there are no significant differences in the number of C.R.'s produced.' Where facts are asserted or denied in this manner in direct contradiction to the actual sources quoted, it would be a task of supererogation to enter into any more detailed argumentation.

It may be worth while, however, to mention one further point. Hamilton discusses the problem of using one-tail tests of significance, and says: 'it would seem that two-tail tests of significance are used almost arbitrarily depending on the mere technicality of hypothesis formation...there would appear to be as much (or as little?) justification for linking by hypothesis conditionability and introversion, hysteria and "Rhathymia", as for linking intelligence and conditionability'. The whole argument suggests an unusual degree of confusion. If a particular postulate generates a theorem, and that theorem is tested experimentally, then the use of a one-tail test for the estimate of significance is obligatory. The general postulate with which we are concerned generates a theorem that normal introverts should condition better than normal extraverts, and that dysthymics should condition better than hysterics. Although not required to achieve adequate significance, the one-tail test is the correct one to use. On the other hand, there is no theory linking intelligence with conditioning, and an overwhelming amount of evidence from the literature suggesting an absence of any relationship. This distinction would seem to be so obvious as not to require mention if it were not for the fact that Hamilton has raised this point.

Apart from conditioning, the only other prediction discussed in detail by Hamilton relates to satiation. I had postulated a functional identity between reactive inhibition and satiation, the similarity of which two functions Köhler himself had drawn attention to, and which had independently been remarked on by Duncan, who discusses in very great detail the remarkable experimental similarities obtaining between these two phenomena. Hamilton's comments are: 'In general it would seem that a theory is being presented that requires a reasonably precise definition of its parameters which would allow the equating of inhibition and satiation, but none is offered...the identity of reactive inhibition, conceived as a variant of cortical conductivity by Eysenck, and satiation as defined by Köhler, remains an assumption for which there is little positive evidence.' Hamilton does not mention the very careful attempt of Duncan, quoted *in extenso* in my book, to establish experimental similarities, and he again contravenes the rules of scientific logic which allow us to formulate any kind of postulate we like provided it generates testable deductions which are then experimentally verified. The evidence here is less clear-cut than in the case of conditioning. Several investigations have verified my original demonstration of the relationship between kinaesthetic figural after-effects and extraversion, although Broadbent has recently reported data suggesting an alternative explanation. Results with respect to visual satiation phenomena have failed to support the prediction. Experimentally the task of testing the theorem is quite a complicated one, particularly as such factors as eye movements interfere in the direct measurement of satiation in the visual sector. We have recently succeeded in eliminating this difficulty and the conditions for performing a proper test of the theorem in the visual area are now available.

Near the end of his paper Hamilton turns to another theorem relating reminiscence effects to extraversion. His treatment of the results there is in many ways typical of his general attitude. The predicted relationship has now been found by a number of different investigators both in this department and elsewhere, and in view of the low *a priori* likelihood of discovering such relationship one might have expected an impartial reviewer to consider this as evidence favourable for the theory under investigation. Hamilton has this to say: 'The relationship postulated... has now been shown to be only very limited and slight in terms of correlations for normal extraverts, while a relationship between neurotic introversion-extraversion and reminiscence has not been found.' For the second part of this evaluation Hamilton refers the reader to a footnote on p. 254 of my book. This footnote does not refer to neurotic introversion-extraversion at all; indeed, neither we nor anybody else have ever carried out reminiscence tests on neurotics of any kind. The footnote does refer to certain time relationships, stressing the point that the expected phenomena are observed after five minutes of practice, but not after 90 seconds. The reader is urged to compare this footnote with Hamilton's statement as a final proof of the total lack of relationship between the writings to which Hamilton refers, and his statements as to what these writings contain. It should also be noted that this error of fact was pointed out to him by letter in connexion with the first draft of his paper, and that nevertheless the same incorrect statement is found in the published article.

It will be noted that I have not taken up all of Hamilton's points. The reason is twofold. In the first place I do not wish to extend this reply to an unreasonable length. In the second place Hamilton's statements bear too little relationship to the facts to make lengthy discussion worthwhile. Furthermore, most of these points have already been discussed in detail in the book itself. The reader who is in any doubt as to whether a given criticism is well taken or not is advised to read Hamilton's presentation of what my colleagues or I said or did in conjunction with what our books and papers show to have been really said and done by us. It will, I think, be found that the relationship is too tenuous to justify any more lengthy reply.

(*Manuscript received* 20 *May* 1958)

NOTE ON BURT'S 'THE DISTRIBUTION OF INTELLIGENCE'*

By J. C. RAVEN

Department of Psychological Research, The Crichton Royal, Dumfries

T. R. Miles has aptly argued that disagreement with Cyril Burt's definition of 'intelligence' as 'innate general cognitive ability' 'is a disagreement on policy'. It is not a disagreement as regards the facts. Test data which have been systematically collected over more than twenty years all point to the fact that the nearer we get to assessing what Burt has found to be the 'general factor' in intelligence tests, the less we can minimize the fact that, for young children, adolescents, adults and old people alike, the usual distribution of test scores does not conform to a Gaussian curve, but is negatively skewed. In my own studies concerned with the development and decline of mental ability with age, this skewness is quite clear, and quite unavoidable. Each attempt to close up the tail of low scores by juggling with the test items, or manipulating the numerical results, has decreased the usefulness of the information obtained both for the practical purpose of differentiating between people and for the psychological study of mental activity in health and illness. Attempts similar to those of Alice Heim to construct qualitatively 'high-grade' intelligence tests, or to rely, as Furneaux does, on differences in the 'speed' of intellectual work in order to spread out high scores, have proved to be limited in their range of usefulness. It would seem that, while many influences can contribute to low-level performance and slow work, those increasing the highest order and greatest speed of intellectual activity a person is capable of are relatively few.

My findings are not only in agreement with the data cited by Cyril Burt; they support most of the inferences he now draws from his data. On the question of policy, however, I disagree with his definition of what is being measured. However certain we are of our facts, I do not think the inferences we draw should be incorporated in the definition we give of what is being tested, or that we need define 'intelligence' at all.

If necessary, we could define 'intelligence' as 'ability to do what an examiner thinks is intelligent', and proceed, as Spearman did, to compare authorities with respect to what they consider 'intelligent behaviour', both as regards the opinions they express and the tests they select. Operational definitions need not be 'circular'. We can observe and assess the behaviour of individuals when thermometers show certain readings, without having to define 'temperature'. Similarly, we can observe and assess a person's behaviour in an environment involving the solution of given problems without having to interpret that behaviour in terms of 'genetic constitution'. Instead of speculating as to its origins in the past, we can make predictions concerning future behaviour, which can be verified by subsequent observation. This is what matters, not only in practice, but from the point of view of scientific theory and method. In view of this, my own policy has been to assess a person's present capacity for coherent perception and orderly judgement, as also his present recall of acquired information and ability for verbal communication, and to ascertain how these two components of a person's present adjustment to his environment are related at convenient age intervals, in health and illness. The results have confirmed the artificiality of fitting living behaviour to a Gaussian curve. (See Raven, 1948, The comparative assessment of intellectual ability. *Brit. J. Psychol.* **39**, 12–19.)

---

* *Brit. J. Psychol.* (1957), **48**, 161.

By comparing people with respect to their present behaviour, we can predict and observe the occurrence or non-occurrence of a particular 'order' of mental activity, and assess a person's speed of work at a given level. We can determine the factors contributing to skilful work or hindering individual adjustment to environmental stresses. We can also assess the rate of maturation and decline of ability with age and health. We can even predict the likelihood of similar maturation in successive generations. This is not merely an art, it is a fruitful line of strictly scientific observation, inference, prediction and verification, in which the definition of what we set out to observe and compare can clarify the inferences we draw from the records we collect. Unfortunately, Cyril Burt's present definition of 'intelligence' as 'the inborn general cognitive factor' compromises the inferences we may legitimately draw by anticipating the conclusions we may or may not ultimately reach, instead of clarifying what we set out to record.

(*Manuscript received* 18 *September* 1957)

THE 'PHENOMENOLOGICAL FALLACY'—A REPLY TO J. R. SMYTHIES

By U. T. PLACE

Institute of Experimental Psychology, University of Oxford

In attempting to demonstrate the fallaciousness of my account of what I have called (Place, 1956) the 'phenomenological fallacy', Smythies (1957) appears to have misunderstood what I meant by that term. He seems to think that on my view to come out with a statement like 'there is a green image in *O*'s mind' is to commit a logical fallacy. Statements by themselves can be true or false, meaningful or meaningless, but only arguments and inferences can be valid or fallacious. To commit a logical fallacy is to draw a conclusion which is not justified or demanded by the premisses of the argument. To say that an inference is fallacious is not to imply that the conclusion is necessarily false, merely that it does not follow from the premisses.

The phenomenological fallacy in my sense is the fallacy of supposing, for example, that the statement '*X* looks green to *O*' commits us logically and inescapably to the conclusion 'there is a green image in *O*'s mind'. This is not to say that the statement 'there is a green image in *O*'s mind' is logically untenable. It may be possible, as Smythies points out, to develop a sort of sense-datum language in which the statement 'there is a green image in *O*'s mind' is used where we should normally say 'so and so looks green to *O*'.

If such a language were generally adopted, there is a sense in which anyone who said '*X* looks green to *O*' would be committed on switching to the sense-datum language to the statement 'there is a green image in *O*'s mind', just as anyone who said 'the sky is blue' would be committed on translating his remarks into French to the statement 'le ciel est bleu'. But it would still be fallacious to deduce from the premisses 'images in the mind are sometimes green', 'brain processes cannot be green' the conclusion 'images cannot be brain processes', since words like 'green' in the sense-datum language would not mean what they mean in our ordinary physical object language to which statements about brain processes belong.

Another possibility which Smythies seems to be hinting at in the latter part of his paper is that, although we cannot legitimately infer the real properties of images in the mind from the apparent properties of things in the environment, we might conceivably have other reasons, perhaps of an experimental nature, for adopting the view that there is something literally green, i.e. green in the sense that grass, corroded copper and the 'Go' signal of the traffic light are green, in *O*'s mind whenever something looks green to him. If such a hypothesis were established beyond reasonable doubt on other grounds, the inference from the apparent properties of things to the real properties of images in the mind would no longer be a fallacy, since it would be supported by the major premiss 'whenever something appears to have a certain property, there is an image in the mind which actually has that property'.

While this possibility must undoubtedly be conceded, it is difficult to see any reason for giving serious consideration to a hypothesis for which there is no empirical evidence, which is inconsistent with the whole trend of scientific thinking and to which no precise meaning has yet been given.

REFERENCES

PLACE, U. T. (1956). Is consciousness a brain process? *Brit. J. Psychol.* **47**, 44–50.
SMYTHIES, J. R. (1957). A note on the fallacy of the 'phenomenological fallacy'. *Brit. J. Psychol.* **48**, 141–4.

(*Manuscript received* 12 *November* 1957)

PUBLICATIONS RECENTLY RECEIVED

Imagination and Thinking: a Psychological Analysis. By P. McKellar. London: Cohen and West. 1957. Pp. xi+219. 21*s.*

'From the senses to the imagination, from the imagination to thought—such is the order of life and of mind.' It is this route which Dr McKellar has chosen to follow, and it may be said at once that he has produced a most interesting book, not the least of whose attractions is that it re-opens a field of observation and experiment deserving much more attention than it has recently gained. Broadly his contention is that what may, with very general denotation, be called imagery, normal and 'abnormal', pervades and plays an important role in mental life at every level and therefore affords the opportunity for tracing genetic connexions back from processes of creative thinking to those simpler functions such as perception which are susceptible to direct experimental attack. Many people may take the view that this is a hackneyed approach: one pursued in the past *ad nauseam* with little positive outcome. It is true that a vast, diffusely connected, literature has accumulated, ranging over mystical experience, hallucinosis spontaneous and induced, types of imagery, the self-revelations of original and creative thinkers and the unconscious sources of inspiration and intuition. This as a whole no one has been able to make very much sense of. Dr McKellar has read widely in this field, but his book is by no means a mere compilation. It is informed by a definite connexion of ideas, and enlivened by observations of his own. In places, perhaps, one may feel that he has drawn his net too widely. In others valuable, relevant, material has been passed over. Towards the end the author's enthusiasm turns didactic to an extent which rather outruns discretion. Nevertheless, as a whole, the book represents an attractive and courageous sally into a field full of difficulties, but one of considerable potentiality.

The book opens with a discussion of imagination, thinking and dreaming. Here McKellar's aim is to suggest that normal, 'respectable' thinking processes have, both in their motivational aspects and in the sources and character of their imaginal material, points of resemblance to the dream which are not commonly acknowledged, and which can legitimately be made a starting-point in their study. Bleuler's useful working distinction between *autistic* and *reality-directed* thought naturally crops up in the course of this discussion, but it is not McKellar's object to break this down. On the contrary, it remains central to a good deal of what he has to say in the rest of the book, and it would have been helpful if he had examined it rather more closely at the outset. It is a distinction which is in some danger of acquiring the impenetrable respectability of an old scandal. Not infrequently it leads to a gratuitous identification of the logical character of R-thought with its character of being in correspondence with objective fact. This arises because A-thought recognizably lacks logical sequence. Logic thus appears constitutive of what is taken to be its opposite—'R-thought'. And 'R-thought', in turn, by definition, corresponds with reality. But in truth the connexion of thought with reality has less to do with the closeness of its logical texture than with a repeated appeal and submission to established fact, whether perceived or stored. Effective scientific thinking, for instance, often lacks correct logical gestures and tends to rely on adventitious opportunity. Thinking, whose main feature is close logic, on the other hand, is apt to be debarred from correspondence with the environment. Either it possesses insufficient redundancy, and a single false element suffices to overthrow the whole, or it plays with concepts too remote and abstract, however easily named, for contact with concrete experience. Bleuler's dichotomy was introduced as a device useful in the description of symptoms. It is far from exhausting the varieties and contrasts of aspect which thinking processes exhibit. To the present reviewer a good deal of what McKellar contends would have been clearer had he made these aspects more explicit at the start—had he considered, for instance, which is the more logical, which is the more autistic and which is the more reality-related—the paranoic's brooding or the scientist's hunch? However, it may fairly be said that in opening up R-thinking to lines of analysis which many people would regard as appropriate (and fair?) solely to A-thinking, he is taking a useful step in the right direction.

In accordance with the conclusion of Chapter I, he next devotes three chapters to an examination of imagery of all kinds, and related experiences. Rightly, he by no means confines his attention to the commoner, more 'normal' types, but adduces data and reports his own inquiries in the wide and fascinating field which includes involuntary manifestations such as hallucinations, hypnogogic imagery and number-forms. The anecdotal, and sometimes 'psychical', treatment of these phenomena in the past has rather brought them to be regarded as a kind of mental junk heap of somewhat unsavoury character.

Such an attitude is no longer appropriate, for, in some connexions, they can now be approached experimentally. The McGill stimulus-deprivation work and the use of intermittent visual stimulation developed by Smythies and others are examples of this possibility. Indeed the best hope of elucidating these phenomena, and thereby gaining some insight into the more general questions to which the author is addressing himself, would seem at present to lie in the study of their physiological conditions. Dr McKellar is perhaps open to some slight criticism here in that he has scarcely come to grips with the considerable amount of material already available, especially in the clinical field. In trying to classify and elucidate the phenomena found in 'normal' people, we can hardly afford to neglect pointers from the field of the neurologically abnormal.

The main types of 'individual subjective experiences' discussed by McKellar are the following: Hypnogogic imagery and hallucination; similar phenomena in the waking state involving hallucination and intrusive, involuntary images; *déjà-vu*; synaesthesia; diagram- and number-forms; and what he calls 'the falling experience'. Wisely he has collected, by interview and questionnaire, a quantity of material about these topics on his own, and although the study of this does not lead to any very definite conclusions, it brings out some interesting points capable of further development. We might instance the peculiar perseverative images which occur in many people after a period of activity involving attention to a visual scene of high contrast. These are involuntary and intrusive; though fleeting, they possess vivid hallucinatory clarity and definite location in the visual field, and they commonly follow such activities as playing chess or weeding a garden bed. They clearly offer opportunity for experimental study, for the conditions which give rise to them could be reproduced in the laboratory. Such a chance is all too uncommon in this field generally.

But if in a number of respects McKellar's treatment of images, hallucinations and related phenomena is stimulating, in others it is somewhat disappointing. This is because his approach to the material is too much that of a collector of individual specimens. Like an older school of botanists, his interest seems to be largely at the descriptive level, and does not extend to ecological considerations. To take, for example, the field of hypnogogic hallucinations; these occur in the same individual in varying combinations, and may show changing trends of emphasis on one type or another over periods of time. They may be associated with marked motor phenomena ('night-jerks'), with paraesthesiae of the extremities and with autonomic manifestations. The hallucinatory elements may include visual (formed and unformed), auditory and olfactory phenomena. In some cases at least there is a suggestion of connexion with migrainous disorder as, indeed, William Gowers pointed out half a century back. All this points to the importance of a more thorough-going clinical study than has yet been made. The relevance of all these things to problems of thought-processes is unlikely to emerge fully until this has been done. Dr McKellar would certainly render a most valuable service by carrying his work beyond its present preliminary stages along these lines.

But it may perhaps be asked what *kind* of relevance the phenomena in question may be expected to have in regard to normal thinking processes? About this McKellar is not as explicit as he might be. He notes, following Galton, 'the occurrence of these subjective experiences among eminent thinkers...'. If it were established that such experiences are *more* common among 'eminent' thinkers than among the ordinary run of mankind, the chief importance of this might well be that it indicates a tendency to instability, perhaps essential to original thinking, in the underlying neural mechanisms. In this connexion it is worthwhile pursuing the line of thought opened up by Sir Frederic Bartlett in his recent book. Here he canvasses the idea that thinking is comparable to bodily skill. The emergence of the brilliant unforeseen thought is like the making of an unexpected master-stroke at tennis. Both come out of 'the blue', sometimes to the surprise of the executant. Now in both cases the fact that the product, though not an outcome of explicit conscious activity, is a well-adapted, appropriate one has implications regarding the underlying mechanisms. Fortunately, in the one case, that of bodily skill, we have the means of examining these implications in terms of conceptions developed by Hughlings Jackson, Sherrington and Head, though Bartlett himself does not explicitly go so far. We have to view the skilled act as the product of a set of motor mechanisms, organized in a hierarchical fashion in which each level makes its contribution, regulated and guided by those successively above it. The forms of activity characteristic of each level may be experimentally examined by the classical methods of neurophysiology: clinically by observation of 'release' phenomena occasioned by disease of higher centres. Adequate understanding of the neural processes at the highest levels required in bodily skills is obviously at present sketchy. But in the attempt to reach this understanding it is essential to bear in mind that these higher level processes do not take place in isolation from the rest of the hierarchy of which they form part. Now if this is true of skills, how much more so is it not true of thought? Ought we not to inquire, by every available means, clinical and experimental, into the underlying levels of mechanism which though essential are normally hidden? Their workings may be glimpsed when in crude and sometimes bizarre manifestation, they for instance occupy the hypnogogic stage or persistently intrude on waking consciousness as automatic image-stereotypes. Dr McKellar has made a start which can claim the support of a sound biological tradition.

One way in which we can feel our way (after a fashion) experimentally in this field is clearly by resort to pharmacological expedients, and McKellar next gives some account of current work, including some of his own using mescaline. The contemporary vogue of mescalinophagy, if it has done nothing else, has emphasized some of the difficulties in the use of such methods. The dramatic character of the phenomena inexpensively obtainable from some of these drugs has perhaps tended to obscure the great variety of individual reactions. And Dr McKellar will probably agree that the attempt to pin down the phenomena by any form of controlled psychological experiment is apt to prove frustrating in practice. However, there have recently been indications that the chemist is gaining some insight into the *modus operandi* of some of these substances and is preparing fresh ones whose chemical structure may bear some known relation to their metabolic effects. It is up to the psychologist to devise experimental techniques, as opposed to observational and introspective habits, to meet the increased opportunities for work in this field which are certainly going to arise.

From artificially induced psychotic phenomena, McKellar passes to some brief consideration of the characteristic features of thought-processes in real psychoses. Although he has some interesting things to say in this connexion, it is difficult to avoid some concern as to the value of generalizing from a bird's eye view of these conditions. Seen in more detail, there are doubtless useful conclusions to be drawn by the experienced observer armed with a fresh set of conceptions. But in Dr McKellar's short treatment the issues important to further inquiry do not tend to emerge in a very clear-cut form. This is manifest, for instance, in his handling of the distinction between the *abstract* and the *concrete*. Certainly this distinction, as introduced by Goldstein in connexion with the brain-injured, and extended by Cameron and others to the psychotic, has, like Bleuler's discussed above, its usefulness in the denotation of certain symptoms. But there is some danger in the course of general discussion such as McKellar's—and he does not seem entirely to have escaped it—of confusing this dichotomy with other features of the connotation of these two terms. In another sense, for instance, it is quite customary and legitimate to draw attention to the abstractness of, say, a neurotic's thought as evidencing its lack of potential contact with adaptive reality—its lack of 'down-to earthness'. The word 'concrete' in Goldstein's use implies failure of adaptation: in another context it may be used to suggest that kind of explicitness without which neither effective logic nor factual relationship is possible. As with Bleuler's distinction, there is evidence here of muddled conception lurking behind time-honoured words. Dr McKellar might have taken the opportunity of uncovering and eradicating it.

In the chapters that follow, we are taken on a tour, perhaps a shade carefully conducted, of fields such as creative thought and activity, the supernatural and the mystical. Better justice might have been done to the latter two topics by a more marked independence of treatment. The interest of both in this connexion is unquestionable, but the particular psychological questions they raise are not entirely co-extensive. However, it is an enviable achievement to have covered, without tedium, ground already so freely trodden by others. Nevertheless, especially in relation to 'creative activity', one is tempted to wonder whether the value of this kind of analysis is not becoming marginal. In particular cases an intellectually sensitive and scholarly biographer or critic, working on fresh material, can still contribute, if obliquely, to our understanding of these things. But the attempt to reach general conclusions of psychological import by re-distillation of such accounts is apt not to rise above the level of intelligent, but infertile, comment.

When he comes, in his penultimate chapter, to scientific thinking we seem to catch here and there the glint of some axe McKellar is gently grinding—for the improvement of the reader, we half suspect. Occasionally there are lapses into unedifying pomposity. On page 168, for example, were find: 'Science, therefore, in its characteristic concern with verification and falsification represents the mature rather than the developmentally earlier form of thinking.' One cannot help wondering, perhaps a trifle unkindly, what Dr Thouless would have made of this statement had it swum into his ken when he was writing *Straight and Crooked Thinking*. The immediate support McKellar offers for it is Piaget's finding that it is only after the age of seven or eight that children are concerned with the truth or falsity of the explanations they frame. If science consists in finding explanations and being concerned about their truth, then the statement boils down to 'Young children don't do science'. But McKellar does not mean to waste our time with anything so tritely true. The chief nigger in the woodpile (though there are others!) is clearly the word 'mature'—a word it might be wiser to banish for ever from all psychological, social, literary and artistic discussion, except in descriptive use within the agreed context of a specific developmental process. In this case, however, what starts with a purely genetic or chronological sense of 'maturity' draws on further connotation to contrive a promise of approval and self-esteem for the reader who will be a good boy and do science. Such verbal conjuring tricks are liable in the long run to defeat even their own ends. A man of Dr McKellar's perspicacity hardly needs reminding that further overtones of that currently overloaded word 'mature' exist which are detrimental to, or incompatible with, scientific or any other form of constructive thinking. But to take exception to small outcrops of this kind is perhaps churlish. On the whole McKellar makes good use of skilfully controlled discursiveness to bear out, in terms of difficult material, his main theme. The psychological process of scientific thought has essentially its autistic and its logical aspects, while another of its vital characteristics is the constant appeal to factual check. With this the present reviewer is in full agreement. We should part company, however, were McKellar to suggest that any particular worthiness attaches to R-thinking in itself.

To some extent, perhaps, this is 'the kind of book which people who like this kind of book will like'. If so, the present reviewer is happy to find himself among such people. But McKellar will certainly succeed in doing more than give pleasure to a class of readers; his work has enough substance and enough backbone to make different appeals to different individuals. To the present reviewer he has rendered the service of laying out afresh an extraordinarily heterogeneous and puzzling field in such a way that it seems clearer than it was before just where points of profitable attack lie, and where, on the other hand, discussion and cogitation, in terms of concepts available at present, are likely to prove nugatory. So much can hardly be said by any reader of many psychological books that appear nowadays.

R. C. OLDFIELD

Ageing and Human Skill. By A. T. WELFORD. London: Oxford University Press, for the trustees of the Nuffield Foundation. 1958. Pp. vii+300. 25*s.*

There is a pregnant sentence in the preface to this book which reads in part, 'The topics have been dictated by the areas in which it has been possible to make studies....' No one without experience of research in the field of ageing can have more than a slight appreciation of the difficulties faced by Welford and his colleagues; perhaps those who have attempted rigorous work in industry come nearest. For this reason much of the criticism appropriate for reports of work done with school children, university students, Services personnel, hospital inmates or any other 'captive' population within a limited age-range, is simply out of place when considering this book. More than that, these studies of age-differences should serve as a much-needed check on the generality of findings obtained exclusively from such sources.

In three chapters—the first two and the last—the reader may find all he needs to set the stage for an appreciation of what lies between. At the beginning the author discusses lucidly some of the probable consequences with ageing of the interaction between organic changes and the material and social environment, and presents the conception of skilled performance from which the vast majority of the Unit's work derived. In the last few pages of the book attention is drawn to the important attitudinal qualities which have *not* been investigated, but which Welford feels 'must surely form a part of the programme of future research in the field of ageing'.

The book also includes a chapter on 'Methods of Studying Age Changes', reprinted by permission from the third volume of the *Ciba Colloquia on Ageing*. Strictly speaking, this title should be lengthened by the words '...and Age Differences', since hardly any studies of *changes* with age have yet been carried out. Welford of course makes this clear and also emphasizes it later: 'Research aimed at studing the results of long experience would seem inevitably to involve longitudinal studies following individuals over a considerable period of years, and imply the need for the highly stable investigating groups that these longitudinal studies require.'

In this chapter on method the author draws attention to the need for samples 'in which older and younger people are *comparable*' (his italics), and goes on to suggest seven methods for partly ensuring that this need is met. Several of these are at present almost impossible of execution, but others seem feasible even though difficult. For example, one can choose two or more groups of widely different background in each age-range, or one can use preliminary testing—with a vocabulary test, for example—to select the groups. Unfortunately, practically no information on this is given when reporting the experimental work forming the bulk of the book. The reader is therefore unable to feel at all certain about the extent to which data obtained from samples of different ages can be compared. It is also true that, probably in the interests of detail-reduction in the text, the actual *numbers* of subjects are not usually stated. These points of course have more cogency when the experiment being reported is not published in full anywhere else, but it would have been more satisfactory had the author made a practice of reporting the numbers and general background of all age-group samples, and indicated what method of control, if any, had been possible in each experiment.

The experiments described in the book are frequently extremely ingenious, and the methods used to obtain a maximum of information from each experiment impressive. Welford and his colleagues were clearly right in distrusting overall 'scores' of performance, and have set a standard by which work of this kind should in future be judged. On the other hand, while it is certainly true that industrial studies can serve as a check on experimental data obtained in short work sessions, one may still entertain some doubts on this score, and seek more laboratory studies of the prolonged miniature work situation type—if suitable and willing subjects can be found throughout the age range!

According to a list provided in the book, the Nuffield Unit during its ten years of life had twenty-four members, for periods ranging from 1 to 10 years. The work reported in this book represents 91 man-years of research, and is thus probably the greatest single investment ever made in the study of ageing. No unprejudiced reader of this book can fail to recognize the worthwhileness of this effort; no one who understands the difficulties of team-research will fail to perceive the unifying result of drawing almost all one's colleagues from the same source. Everyone who read the earlier and very preliminary report by the same author will applaud the striking advance in content and presentation now achieved: this is compulsory reading for all students of ageing, and nearly so for all research psychologists whatever their field of interest.

ALASTAIR HERON

Perception and Communication. By D. E. BROADBENT. London: Pergamon. 1958. Pp. v+338. 55*s*.

The twelve chapters are headed: *Hearing and Behaviour, Selective Listening to Speech, Verbal and Bodily Response, The Assessment of Communication Channels for Ease of Listening, The Effects of Noise on Behaviour, The General Nature of Vigilance, Some Data on Individual Differences, The Nature of Extinction, Immediate Memory and the Shifting of Attention, The Selective Nature of Learning, Recent Views on Skill, Retrospect and Prospect.* These have been listed in full because the title of the book may not adequately prepare the reader for its contents. He may be surprised at the absence of most of the traditional topics of visual perception, at the small amount of space given to communication in the ordinary sense of the word, and at the prominence given to learning, memory, and skill.

Throughout the book the author is concerned with applying some of the concepts of communications engineering to behaviour, but apart from the degree of unification this provides the book does not hold together, chapters tending to deal with their separate topics in the manner of review articles. The theme recurring most frequently is that the perceptual system can be regarded as a communication channel of limited capacity preceded by a filter. The primary evidence comes from experiments on selective listening. When the channel is occupied by one set of signals, other signals may be systematically discarded or filtered out. Experiments on immediate memory indicate that there is a short-term store in which incoming signals may be held. On an information flow diagram this store can be located between the senses and the perceptual channel. The maximum storage time is of the order of seconds, but signals can be kept in circulation by being passed through the perceptual system and fed back to the store. To account for the relatively permanent effects of learning, conditional probabilities of past perceptual sequences are stored separately on the output side of the perceptual channel. This permanent store and the perceptual channel jointly lead to a system for varying output (behaviour) until some input (perceptual event) is obtained. No attempt is made to use information theory quantitatively.

Those who dislike 'models' and information theorizing should not be put off. A large part of the book is concerned with close and careful analysis of experimental data. On the theoretical side the author states his position without being doctrinaire or dogmatic, considering other points of view with sympathy.

The best parts of the book are those that deal with areas in which Broadbent has established research interests—multi-channel listening, the effects of noise, vigilance, immediate memory. They include excellent summaries of experiments and critiques of theories. Less balanced, more speculative and tenuous are the discussions of the permanent storage and output side of the perceptual channel, based on theories of Uttley and Deutsch. The chapter on individual differences is mostly limited to introversion-extraversion, but it seems more fitting to congratulate the author on including individual differences of any sort within his theoretical framework than to cavil with the chapter's limitations.

Broadbent contrasts the concepts he used with those of S-R psychology. There can be no doubt that his point of view is fruitful, but the antithesis may not be fully justified. Expectancy theory is said to have been neglected because of the vague and 'anthropomorphic' associations of words like 'cognition'.

Regarding his own point of view, he believes that since its main features have already been embodied in mechanical models there can be nothing mystical about it. There are two dangers here. The first is the risk of supposing that the construction of a model does anything to validate a theory. A well-known theory of instincts could be 'actualized' by any plumber, but this would not contribute to its acceptability. The second is the risk that words like 'filter' have such a respectable scientific ring to them that they will be used loosely. When Broadbent's filter is said to 'choose' and 'recognize' we may well be suspicious. The 'filter theory' that appears in this book makes only a start at accounting for the selectivity of perception. While the action of a filter can be precisely specified in Broadbent's experiments it remains to be seen whether the concept will bear expansion without a serious loss of precision. One suspects it may suffer a fate like 'homeostasis', and an extension of its usage may make it as difficult to envisage physiologically as a superego.

The book ends with a discussion of strategies for research. Experiments should be designed to maximize information. Broadbent disapproves of hypothetico-deductive methods and over-eagerness to make detailed and quantitative predictions. Experiments designed to confirm predictions at long odds do not provide the best strategy. While agreeing with this point of view, it remains true that where a behavioural outcome results from the action of several determinants, it may be impossible to make any type of prediction without at least the rough quantification that may be necessary in specifying the relative importance of each factor. This was Hull's point of view.

The reader cannot fail to notice how often British work is cited. A sample of 100 references gave 35 British, one European and the rest American. This will be welcomed for a variety of reasons. Some of the topics that the author has chosen would make the relatively high proportion of British references inevitable, but there are a few other occasions when the parish pump might have been left further behind. The book deserves a better subject index. Diagram 5 is upside down.

There can be no doubt that the book is important and that it will be well received and frequently cited. Its moderate, well-founded claims may make converts among those who have looked askance at information theory. A less heterogeneous slimmer (and cheaper) volume confined to the topics of perceptual span and selectivity might have been more pleasing to University teachers at least, but we should be thankful for what we have been given. It is well worth having.

R. L. REID

Ciba Foundation Symposium on the Neurological Basis of Behaviour. Ed. by G. E. W. Wolstenholme and C. M. O'Connor. London: Churchill. 1958. Pp. xii+400. 52*s*. 6*d*.

This Symposium was held in London to commemorate the birth of Sherrington on 27 November 1957. It does not, however, reveal his influence as strikingly as might have been expected and, apart from Prof. Eccles, his best-known former pupils and collaborators were conspicuous by their absence. None the less, the volume does provide a convenient summary of much recent work on the nervous system in relation to behaviour, with special reference to recent advances in electrophysiological technique. It also provides evidence that, at least in America, behaviour is regarded as a reputable object of physiological inquiry.

The topics covered in this volume are so varied that a short summary is scarcely possible. They include the behaviour of nerve cells (Eccles), intracerebral stimulation in conscious animals (Andersson, Olds), effects of cerebral ablation (Bard, Kluver), the temporal lobes and memory (Penfield), neurophysiology of conditioned reflexes (Gastaut, Jasper), and effects of hormones (Harris), enzymes (Rosenzweig) and drugs (Elkes) on cerebral activity and behaviour. Dr Magoun appears in an unusual role as historian of the brain–mind relation, and Dr Penfield in his more accustomed one as the Gall of our day. The discussion which follows the papers is for the most part pertinent, though largely confined to the technical aspects of the communications. Only Dr Kluver stands out as a man conscious throughout of the broader issues of the Symposium and with an encyclopaedic knowledge of the whole field. Fittingly, he was chosen to wind up the final discussion.

It is a pity that no part of the Symposium was set aside for discussion of the difficulties which arise in the interpretation of stimulation and ablation experiments and, more generally, of how the neurological basis of behaviour is to be envisaged. As a record of current work making use of the newer electrophysiological and pharmacological techniques this volume will be widely welcomed. But as an attempt to exploit the breakthrough achieved by Sherrington fifty years ago in his *Integrative Action of the Nervous System*, the present operation can only be described as a series of unintegrated raids on a narrow front. Perhaps the Ciba Foundation, whose generosity made this Symposium possible, will now promote an operation designed to extract from Prof. Kluver the book we are all waiting for.

O. L. ZANGWILL

Diagnostic Performance Tests: A Manual for Use with Adults. By BORIS SEMEONOFF and ERIC TRIST. London: Tavistock. 1958. Pp. xvi + 176. 32*s*.

It has of late become the fashion to criticize the use of mental tests on the ground that the characteristics with which they deal form only a fraction of the total personality, and that a mere quantitative mark or measurement cannot possibly do justice to the qualitative aspects of the human mind or to the varying ways in which these different aspects are fused and integrated to form a unique individual whole. The earlier investigators, as Prof. Vernon points out in his preface to this book, made no such mistake, and invariably stressed the need for combining psychometric procedures with the clinical or case-study method of approach. But during the last twenty years or so there has been a growing tendency for the two complementary techniques to fall into separate hands. It is therefore doubly welcome to find the authors of this new contribution to diagnostic testing explicitly insisting on the need for both. 'A testing session', they maintain, 'should be regarded as a social situation involving the interaction of personalities: as such it can yield insight far transcending the result of a mere mechanical examination or a comparison of test scores.' And accordingly, as they explain in their preface, 'the approach to psychological testing which this book exemplifies is one that attempts to bring together both qualitative and quantitative methods in a rigorous yet flexible manner'.

The book itself is intended to serve as a practical manual rather than a systematic or theoretical treatise. Its primary purpose is to furnish instructions for the use of a set of five performance tests, some entirely new, others based on well-established procedures, all carefully standardized and combined in a battery for assessing the psychological efficiency of adults. Most of the work on which it is based was carried out during the recent war, first at Mill Hill Emergency Hospital, and then at the War Office Selection Boards. These Boards, set up in 1942, were concerned primarily with selection of Army officers; and several of the tests here described were then employed as 'confirmatory techniques' to secure a more trustworthy rating of the intelligence of the candidates who were interviewed. Later on, at a special Assessment Board of the Inter-Services Research Bureau at 'Pemberley' (a code-name borrowed from Dr Morgan's account in *Spies and Saboteurs*), the tests were put to still wider uses: here the 'performance test situation' was treated as an opportunity for making qualitative observations, and so eventually securing an all-round assessment which would, so far as possible, take into account both cognitive and non-cognitive factors in estimating each candidate's personality.

In the actual conduct of the work, Mr Trist, who had served as Research Psychologist at Mill Hill Emergency Hospital, was the pioneer. New assignments soon afterwards cropped up for which an extensive use of performance tests became essential; and for this Dr Semeonoff became responsible. He has since continued the work at the University of Edinburgh, and thus collected additional data. At the request of the Tavistock Institute of Human Relations, which provided a grant in aid, both the war-time and the post-war material have now been brought together to form the present manual. For the general substance of the book and for the views expressed, the credit belongs jointly to both authors; but the text itself has been the work of Dr Semeonoff. As Prof. Vernon observes, it is rare to find two individuals who are able to appreciate and to combine both quantitative and qualitative methods of approach.

The main battery consists of five basic tests: (i) the Semeonoff-Vigotsky Test (a test of concept-formation in which the examinee is required to classify a set of coloured blocks of various sizes and shapes into the most appropriate categories); (ii) the Trist-Hargreaves Test (in which the examinee is required to classify and reclassify a set of twelve coloured pieces successively into four, two, and three groups in various ways, and then to make pairs corresponding to certain specimen pairs); (iii) the Trist-Misselbrook-Kohs (a new version of the well-known Kohs Block-Designs Test, which has been modified so as to throw light more specifically on the examinee's capacity to learn); (iv) the Carl Hollow Square (a series of graded problems carried out with a form-board and twenty-nine separate pieces, with both straight and bevelled edges); and (v) the Revised Passalong Test (a modification of Alexander's Passalong test adapted for testing adults rather than children). One chapter is devoted to each test in turn; and describes in detail the apparatus required (now obtainable from the National Foundation for Educational Research), the instructions for applying the tests, the procedures for scoring and evaluating the results, and supplementary suggestions for further research. A series of excellent photographs, diagrams, and coloured plates illustrates the type of material used and the mode of administration and assessment. A number of supplementary tests—the Mitchell Vocabulary Test, a Shortened Wechsler Verbal Scale, and translated versions of the Reasoning Tests—together with illustrative case-material are described in the appendices.

For reasons of security the relevant dossiers and other documents relating to the subsequent histories of the individuals tested have not been available. Hence it has not been possible to present first-hand data demonstrating the validity of the assessments thus obtained or of the interpretations that might be suggested for the different types of behaviour observed during the actual testing. However, the two concluding chapters give detailed tables of norms and a summary of the chief statistical data so far available: these consist of correlations between the tests, brief factorial studies, and a note on the reliability of the method devised for converting the raw test scores into 'intelligence gradings'.

The whole book is based on a vast amount of accumulated experience and represents an important advance in the technique of psychological assessment. It forms a most valuable addition to the psychologist's equipment for exploring adult personality and assessing adult intelligence; and can be warmly recommended to psychologists, psychiatrists, and all who are interested in the problems of individual diagnosis.

CYRIL BURT

Information, Decision and Action: A Study of the Psychological Determinants of Changes in Farming Techniques. By F. E. EMERY and O. A. OESER. Melbourne University Press. London: Cambridge University Press. 1958. Pp. xiii+132. 21*s*.

This book reports a study of communication between scientist and farmer, carried out by members of the Psychology Department of Melbourne University. Australian scientists and agricultural administrators had felt for some time that farmers were slow to use the results of research, and a study of the working of the advisory services, administered, as in Britain, by District Advisory Officers (D.A.O.'s) was therefore undertaken.

A small sample of graziers (sheep-farmers) was intensively interviewed about their farming practices and about the channels by which they received information. The investigators developed several indices which appeared to be related to exposure to new information, and hence to the adoption of progressive practices. (1) An index of urbanization (involving an instrumental rather than a traditional attitude to knowledge, and related to post primary education and a history of urban employment). (2) An index of situational motivation (based on age, health, wife's co-operation, allowing sons to share responsibility). (3) An index of conceptual skills and managerial capacity (using credit facilities, planning ahead, attending to maintenance).

These independent determinants were expressed as trait indices, and correlated with various sets of progressive practices. Certain groups of practices, such as pasture improvement, were shown to yield scalable universes of Guttman type.

The main channels of communication were (*a*) mass media, i.e. radio and the printed word, and (*b*) face-to-face contacts with the D.A.O.'s, with commercial representatives, with other farmers, and field days and demonstrations.

The findings and recommendations cannot adequately be set out in a short review. It was found, however, that the larger farmers contacted the D.A.O.'s more often than the small farmers. This finding was attributed to status and educational differences between D.A.O.'s and small farmers. Personal contacts with the D.A.O. were more likely to induce farmers to adopt new practices than field days. The large-scale progressive farmers were themselves 'influencers' who could in some ways communicate with other farmers more readily than the D.A.O.'s. It was therefore recommended that agricultural extension work should concentrate on using these progressive farmers as agents of transmission.

The data are most methodically and imaginatively analysed, but the generality of the causal inferences made is rather limited by the small scale of the study. Although statistically significant, it seems likely that many of the observed relationships would change if the investigations were repeated in another area.

The theoretical background to the study was the theory of communication developed by Heider and by Newcomb. The intensive correlational analysis of interview material of the kind used here seems strangely neglected, both by psychologists and by sociologists in this country; this book could certainly be read with profit by all undertaking applied social research in new fields.

H. MADDOX

The Psychology of Interpersonal Relations. By F. HEIDER. New York: Wiley. London: Chapman and Hall. 1958. Pp. ix + 322. 50*s*.

Person Perception and Interpersonal Behavior. Ed. by R. TAGIURI and L. PETRULLO. Stanford University Press. London: Oxford University Press. 1958. Pp. xx + 390. 60*s*.

These two books are a welcome sign that at least some psychologists are now turning to the study of questions recognized as psychologically significant in everyday life. For example, how do we form impressions of other people? Why do we often take an immediate liking or disliking for them? Do we 'perceive' persons in the same way as we perceive things? What is the basis of our feeling that someone deserves to be rewarded or punished? How exactly do we induce someone to do what we want?

Prof. Heider is not much concerned with presenting empirical data. He assumes that a great deal of valid knowledge about interpersonal relations is already familiar to us from everyday experience and the task he sets himself is one of clarifying this knowledge and of 'conceptualizing' it in systematic fashion. In this venture he is influenced by Lewin's method of topological analysis, to which he pays tribute, but which he considers inadequate for representing the two-person situation in which he is primarily interested.

He proceeds by attempting to construct a language that will allow him 'to represent, if not all, at any rate a great number of interpersonal relations, discriminated by conventional language, in such a way that their place in a general system will become clearer'. Beginning with words that seem to be vaguely related to one another, he then seeks for underlying concepts that make it possible to represent the interrelations systematically. The fundamental concepts he employs—life-space, perceiving, causing, can, trying, wanting, suffering, sentiments, belonging, ought—are a sort of psychological 'basic English'. A notation for these concepts, explained in an Appendix, is used throughout the book. Whether psychologists, already confused by a Babel of Tongues, will be prepared to acquire yet another meta-language is a moot question. How many would wish to learn to use, for instance, '$\hat{x}(pW:o$ not Perc $x)$' to mean 'all the x's that p wants o not to perceive'?

In so far as Prof. Heider's work rests on verbal analysis, it may be open to the objection that one cannot build a universally valid psychological system in terms of expressions encountered in the English language. This remark should not however be taken to detract from the great interest which this original work, the outcome of many years of study and thought, will doubtless have for a considerable number of readers.

The second volume consists of contributions to a Harvard symposium held in 1957. Important articles by R. B. McLeod and Fritz Heider, previously published elsewhere, are also included. As is usual in the case of collections of this kind, the papers vary widely in quality, and in spite of the editors' commendably strenuous efforts, the volume as a whole can hardly be said to leave the reader with an impression that very much has been achieved or indeed that the volume was worthy of publication in its present form. Several chapters seem to be little more than 'thinking aloud'. The heavy jargon in which parts of the book are written will be repellent to many readers. It is to be hoped that the neologism 'perceivee' introduced by one contributor will not become popular.

One of the more instructive papers is contributed by W. H. Ittelson and C. W. Slack on 'The Perception of Persons as Visual Objects', in which the well-known Holway–Boring study of size-constancy is taken as a paradigm. Their study seems to come in sight of an experimental approach to limited aspects of the subject-matter of the book. The experiments described by them relate to: (i) the apparent distortion of the metric properties of persons due to the observer's use of aniseikonic glasses; (ii) Engel's novel use of the phenomenon of binocular rivalry; and (iii) size-distance relationships in the perception of persons. Ingenious as is their effort, it only touches the fringes of the phenomena of social perception.

Another interesting paper comes from J. S. Bruner, D. Shapiro and R. Tagiuri on 'The Meaning of Traits in Isolation and Combination'. This takes an earlier report of Bruner ('Going beyond the Information') a step further. The authors' aim was 'to investigate the principles that relate the inferences drawn from single traits to the inferences made from combinations of the same traits', and several suggestions emerge which may be of promise in future research. In a separate paper on 'Social Preference and its Perception', R. Tagiuri addresses himself to the problem of the resemblances between individuals in perceiving 'regularities' in others. An allied topic is considered by W. L. Hays in a chapter on 'Trait Implication and Trait Similarity' in which he sketches two models for describing the possible formal relations among perceived traits which might underlie inferences made from one trait to another.

Also of interest is a study reported by P. F. Secord on the part played by physiognomic cues in forming impressions of personality. The author attempts to identify the factors which give 'direction' to the impressions which are formed on seeing a photograph. He distinguishes three types of factor: cultural, inferential (e.g. a high forehead signifies intelligence, thin lips mean asexuality), and attributes of the perceiver.

A most timely contribution comes from Joan Creswell in a perspicacious critique of the ideology of the 'scientific' psychologist, plodding away in his laboratory and happily cultivating the fantasy that he is advancing science.

In a brief review of a composite book of this character it is not possible to do justice to all the contributors. Failure to refer to them must not be regarded as a slight but is due to the limitations of space imposed on the reviewer.

JOHN COHEN

Readings in Perception. By DAVID C. BEARDSLEE and MICHAEL WERTHEIMER. Princeton: Van Nostrand. 1958. Pp. xiii+751. 66*s*.

This collection of fifty-two readings is one of a new series of books under the general editorship of D. C. McClelland. My first impression was one of pleasure when I noticed three recent and interesting papers: Miller on the magical number seven, Bruner on perceptual readiness, and Hochberg on the Cornell symposium on perception. The readings are grouped into five main categories: Biological Factors in Perception, Psychological Measurement, Perception of Objects and Events, Imagery and Fantasy, Recognition and Identification of Figures. The third category is much the largest and has thirty-eight papers to itself. The selection of papers must necessarily be a difficult task, but I could have wished that there were not so many topics in this section which seem to my possibly prejudiced eye to be rather associated with Gestalt theorizing. I also cannot quite see the necessity for three papers on sound localization.

The physiological basis of perception is poorly represented. The only sizeable parts on this topic are Morgan's article from the Blake and Ramsey symposium, a study of the effects of adrenochrome, a little of Köhler on his own theory, the various sections of Bruner's paper and his description of the recent work of Hernandez-Péon *et al.* on the effects of distraction on the electrical activity of the cochlear nucleus of cats. I should have thought that room could have been found for a whole paper by or on Hebb. This points up to what seems to me to be the main defect of the collection: there are very few really controversial papers. Surely there was space for a detailed critique of Köhler. The paper by Lashley, Chow and Semmes would have been ideal, as it also contains a valuable summary of some of this work.

These readings are intended to supplement or to serve in lieu of a one- or two-semester course in perception, but I am afraid that the student reading this without assistance would need much more guidance. The introduction is two and a half pages in length and the only additional help is the occasional footnote bringing the references of some of the papers more up-to-date. Three of the papers are from the *Scientific American*. The references which are at the back of this magazine should have been included rather than leaving the student with nothing to follow up these topics. Hallowell's article is without the original additional and interesting notes. Several of the papers have appeared in other collections and I am beginning to feel that I see Atkinson and McClelland's paper on submarine trainee torture wherever I look. This is the second occasion on which I have seen LeShan's article on time orientation in a collection of readings. Have the editors really studied it? I have found it to be an excellent example for my students on how not to do an experiment. Surely it could have been replaced by something by Cohen.

There are so many papers which could have happily replaced some of those included. What about Murphy on 'Affect and Perceptual Learning', Prentice on 'Functionalism in Perception' and Vernon on the 'Functions of Schemata in Perceiving'? I think that the insularity of the editors is particularly noteworthy. Apart from the early Gestalt work and the extract from Michotte, there are not more than two other papers from Europe. There is much recent work from the Institute of Psychiatry and from the Applied Psychology Research Unit at Cambridge which could have been included.

On the positive side, a collection of so many papers in one binding is sure to be useful. There are several papers which are not easily accessible to the student and a number of new translations. Over half of the papers have been written since 1950, and the three by Bruner, Miller and Hochberg will be very welcome.

BERNARD R. SINGER

Tonality as a basis for the study of Musical Talent. By ERIK FRANKLIN. Göteborg: Gumpert. 1956. Pp. 193. No price given.

Emotion and Meaning in Music. By LEONARD B. MEYER. Chicago: University Press; Cambridge University Press. 1957. Pp. x+307. 43*s.* 6*d.*

The Social Psychology of Music. By PAUL R. FARNSWORTH. New York: Dryden Press. 1958. Pp. xiv+304. $4.50.

These three books on the psychology of music are very different from one another in almost every respect, including content, approach, author's background, and—although one hesitates to say so—readability.

Franklin's book is essentially a research report. It is translated from the Swedish, and occasionally the author's meaning is not quite clear in every detail; furthermore, quotations from authors who wrote in German, or even Dutch, are reproduced in the original language, without translation. As implied in his title, Franklin's main concern is with the experience of *tonality* (i.e. melodic structure as determined by the key in which a tune is written). He uses this experience as the basis of a test of 'tonal musical talent' ('TMT'), in which the subject has to complete an unfinished melody. The results, obtained mostly with musically trained subjects, were factor-analysed in a battery along with other tests, including existing tests of musical ability, and Franklin concludes that musical talent is to be regarded as functioning on two levels. In this he agrees with the early work of Révész and with the results of analyses by British writers in the same field (Wing, McLeish). Franklin describes his two 'talents' as sensory or 'mechanical-acoustic', and 'judicious-musical', which latter, he says, 'places the ability to discriminate at the service of music'.

In contrast, Meyer's book is the work of a philosopher and composer. Unlike Franklin, he makes little direct reference to previous writers in the psychology of music. Nevertheless, his starting-point is a psychological one: of the two opening chapters, the first (headed simply 'Theory') develops the author's views—deriving largely from Hebb—on the nature of emotion; the second is entitled 'Expectation and Learning'. The main argument may perhaps be summarized in the author's own words. 'Affect', he says, 'is aroused when an expectation—a tendency to respond—activated by the musical stimulus situation, is temporarily inhibited or permanently blocked'. Then, after showing how 'deviations can be regarded as emotional or affective stimuli', he concludes that 'once the norms of a style have been ascertained, the study and analysis of the affective content of a particular work in that style can be made without continual reference to the responses of the listener or critic. That is, subjective content can be discussed objectively.'

The greater part of the book consists of an elaboration of this and cognate theses. Over a hundred examples in musical notation are used as illustrations, and unless one is able to read these (in the proper sense) much of the point must inevitably be lost. But even so it is a powerful and provocative book, if somewhat heavy going.

Most of Farnsworth's book (which, incidentally, is exceptionally attractively produced) can be read more easily than either Franklin's or Meyer's. The term 'Social Psychology' is here used to cover almost everything: the history and structure of musical scales; music as communication; musical taste; the nature and testing of musical ability; applications in therapy and in industry. All sections are fully documented, and the book is one to which one can look forward to returning for reference. Some sections are, inevitably, a little disappointing; thus, the chapters on taste amount to little more than an enumeration of the results of popularity polls, counts of encyclopedia coverage, and the like, the only moral of which seems to be that a certain—if limited—degree of consistency is found between lists obtained under differing conditions. In general, the book appears perhaps to better advantage in its factual than its interpretative passages, and frequently one finds oneself pencilling question marks in the margin. That this should be so emphasizes how imperfectly the real emotive quality of music, particularly for the untrained listener, is as yet understood.

B. SEMEONOFF

Structuren der Rundfunkprogramm-Auswahl bei Hausfrauen. By OSKAR GRAEFE. Münster: Aschendorff. 1958. Pp. viii+172. Paper DM 17.75; cloth DM 19.25.

The routine part of audience research consists in ascertaining, at frequent intervals from representative samples, how many have listened or viewed a given programme and, of those, how many have liked, felt indifferent or disliked it. The listening or viewing figures and the popularity ratings tend to be presented for the population as a whole or else for given sections of the community. Such information is of course important for the planning of programmes; there is, however, something distressingly piecemeal about it from the point of view of the social psychologist. It is therefore refreshing to come across an examination of radio programme preferences which goes beyond such an approach and attempts to determine, first, whether preferences for given programmes tend to go together and, secondly, whether one can isolate given types of radio listeners. Graefe took a quota sample of German housewives, all of whom had access to the Nordwest Deutscher Rundfunk and asked them to indicate for some 400 programme titles in turn whether they would or would not wish to listen to them (the information was sought on a five-point scale). Graefe carried out two factorial analyses of the responses: in the first, he isolated four factors of programme preferences ranging from a factor of preference for intellectually stimulating and demanding programmes to a factor of preferences for programmes where the reaction sought is an emotional one, preferably involving excitement and suspense. Graefe then took the analysis one step further and, using Stephenson's inverted factor technique, isolated six different types of listener, each being accounted for by a particular set of programme preferences and dislikes. Examples of two such types are the housewives with interest only in matters pertaining to *Küche, Kinder und Kirche*, and those with special interest in information about the outside world.

Having isolated these six types, Graefe shows that they are in turn related, although not always in the expected direction, to education, the occupation of the husband, age, church attendance, political allegiance, the reading of newspapers and periodicals, the amount of radio listening and cinema attendance. One would wish that Graefe had done more with these data; the book develops an interesting approach, one which deserves closer study and might well lead to the isolation of taste patterns which might then in turn be related not only to sociological but also to psychological variables of temperament and personality. The book is written in German in a difficult and somewhat heavy-going manner; it is hoped that a paper on the subject will appear in an English journal.

HILDE HIMMELWEIT

Chemotherapy and the Central Nervous System. By HENRY MCILWAIN. London: Churchill. 1957. Pp. 328. 45*s*.

Pharmakopsychologie und Psychopathologie. By WOLFGANG DE BOOR. Berlin: Springer. 1956. Pp. 291. DM. 39.60.

Psychotropic Drugs. Ed. by S. GARATTINI and V. GHETTI. Amsterdam: Elsevier. 1957. Pp. 606. 95*s*.

These books cover roughly similar ground, but they do this in very different ways.

Chemotherapy and the Central Nervous System is a text-book, and one of the first in this field. It contains chapters on general principles and on the chief groups of drugs used in the treatment of disorders which involve the central nervous system, including mental illness. Prof. McIlwain's own investigations have mainly been into the biochemistry of isolated nervous tissues, and he treats these aspects most fully and with great distinction. He has also included a useful and sober survey of much relevant clinical, pharmacological and psychological data. His book is therefore likely to be of great value to anyone who wants to know on what kinds of evidence the use of drugs in psychiatry and neurology is based and how information derived from different approaches may fit together.

Dr de Boor intended *Pharmakopsychologie und Psychopathologie* to be 'eine Art Grundriss' of the subject, though he seems to have based it almost exclusively on investigations published in German. For example, out of over 100 references on alcohol, four appear to be in English and two in French, and in some chapters the proportion is still smaller. Individual drugs are dealt with one by one, and their effects are usually described in terms of case histories and accounts of 'Selbstversuche'. Experimental evidence, of which a great deal exists by now, is seldom quoted. All this makes the book appear a little conversational and old-fashioned. Nevertheless, Dr de Boor deserves thanks for introducing one to so much unfamiliar and interesting literature and for his descriptions of the 'unantastbaren Eigenschaften' of experiences induced by drugs. Such descriptions are apt to sound particularly fruity of course in German.

Psychotropic Drugs consists of the original papers and reviews which were presented to an international symposium at Milan in May 1957. The term 'psychotropic drugs' seems to have been specially created, and it has been liberally interpreted. Some 160 authors have contributed. The book contains a large amount of interesting material which has been grouped into five sections: (i) Biochemistry of normal and altered cerebral function; (ii) Behavioural effects of psychotropic drugs on animals; (iii) Electrophysiological basis of normal and altered cerebral function; (iv) Pharmacological basis of normal and altered cerebral function; and (v) Clinical experiences. There is a detailed index. For these reasons the information in this book is much more accessible than in many other publications of this kind. Psychologists might be most interested in sections (ii) and (iii) where several papers describe how aspects of behaviour—e.g. sensory input, conditioned avoidance, self-stimulation of the brain—may be altered and analysed by means of drugs. *Psychotropic Drugs* is elegantly printed and illustrated.

HANNAH STEINBERG

A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. By H. B. English and A. C. English. New York, London and Toronto: Longmans Green. 1958. Pp. xiv+594. 60*s*.

In 1934 Dr English served as a member of the Advisory Board for Warren's *Dictionary of Psychology*, an excellent book which has maintained its value over a quarter of a century. During that time, however, a great deal has happened in psychology, and it is an educative as well as an entertaining pastime to compare the present work with Warren's dictionary and find out what is new, and what is considered dated or unimportant.

A main difference in treatment is that in contrast with Warren's use of a somewhat heterogeneous but very useful set of Appendices, Dr and Mrs English have preferred to put everything into a single alphabetical sequence. This has the merit of simplicity, but it might perhaps have been more convenient to have had such things as Hullian and Rorschach symbols and statistical formulas (the authors' preferred form of the word) listed separately in tabular form. A good deal of detailed information from 'related fields' (e.g. anatomy of the central nervous system) has also been sacrificed.

Perhaps the most individual feature of the book is that the authors supplement many of the definitions with comment that is usually illuminating, often pungent and only occasionally irrelevant. In this connexion the reader is invited to look up *xenoglossophilia*, which makes an unexpected appearance at the end of the list of 'General Articles' which contain 'material that has bearing considerably beyond the entry item'. It would be unfair, as the film reviewers would say, to divulge the content of this truly Johnsonian entry, which gives a concise statement of the authors' views on psychological terminology. Having read it, one realizes the significance of the book's subtitle—'*A guide to Usage*'.

B. SEMEONOFF

Theory of Psychoanalytic Technique. By Karl Menninger. London: Imago. 1958. Pp. xiii+206. 25*s*.

This collection of papers is intended primarily for trainee psychoanalysts, but forms a reasonably coherent and readable whole for any British reader who is interested in the rather different direction which the psychoanalytic movement is taking in the United States. Quite apart from Dr Menninger's status, the care which he has taken to compare his views with those of other American practitioners makes it possible to regard a great deal of what he says as representative of their form of Freudian orthodoxy.

For Dr Menninger, psychoanalysis is a contractual relationship between a patient and an analyst, in which the patient 'regresses' into a more or less infantile attitude towards the analyst. He expects more interest, love and advice from the analyst than the analyst does in fact give him; and it is this 'controlled frustration' which is so far the best method discovered 'for allowing the patient to discover his style of, and his conditions for, loving and hating'. Dr Menninger goes on, however, to make it clear that 'interpretation' is still the characteristic psychoanalytic technique for reducing the regression or transference neurosis, which means that he does not favour 'attempts to "handle" the transference by artificially assuming various roles or making artificial manipulations in the conditions of treatment...'. The patient's emotional attitude and behaviour toward the analyst must be 'converted into explicit verbalizations' and be understood by interpretation.

The most interesting and convincing passages are those in which Dr Menninger gives advice on matters of technique. His advice is not very daring: he does not even mention, let alone advocate, any of the attempts which are being made—and especially in his own country—to shorten treatment by hypnotizing, drugging or grouping patients. At one point he is unorthodox enough to say that a patient who spends a lot of time in childhood reminiscences may be indulging in a form of resistance; and this is probably a concession to the late Karen Horney, who was, I believe, the first to assert this. He does, however, make the extremely interesting, and probably original, observation that when a patient is making progress in the consulting-room his flow of thoughts usually follows a recognizable sequence; at first he talks about present-day reality, then about his relations with the analyst, then about his childhood, then about reality, then about the analytic relationship, and so on. Any other sequence is a sign of serious resistance. If true, this is a valuable piece of technical information.

Dr Menninger also has practical advice to offer on the problem of bringing analysis to an end. It is disappointing to find that he has none to offer about the possibility of following-up his patients during the year or two after the end of treatment. This is notoriously a precarious period for the patient, and yet it is very difficult to find anything factual about it in the literature.

On the subject of fees Dr Menninger's views are a little naïve. He gives the impression that it is an essential part of treatment that the patient should pay the analyst's full fees, since otherwise he does not value the treatment or have an adequate incentive to co-operate. The argument may be true enough, but at most it justifies the requirement that the patient pay as much as he can afford. Dr Menninger seems to have jumped to the conclusion that it justifies the analyst in refusing patients who cannot afford his fees: and he even says that one of the circumstances in which treatment may have to be broken off prematurely is when the patient has serious financial problems. I know of at least one clinic in this country which successfully works on the principle of charging the patient according to his means, and meeting the rest of the cost from voluntary subscriptions. It would be interesting to learn from psychoanalysts in the National Health Service what they think about the necessity of fees. NIGEL WALKER

Emotional Factors in Pulmonary Tuberculosis. By DAVID M. KISSEN. London: Tavistock. 1958. Pp. xvii+246. 35*s.*

Most clinicians and not a few psychologists are bemused by the wide range of disorders now labelled psychosomatic. Dr Kissen's book will leave nobody unconvinced of the importance of emotional factors in pulmonary tuberculosis. The results of three controlled investigations are presented in much detail. All these studies show that emotional factors precede the onset of tuberculosis in a significant proportion of cases as compared with controls. Actual or threatened breaks in a love link (romance, engagement or marriage) emerge as the most significant trauma preceding the onset or relapse. The outstanding personality trait found in 100% of tuberculosis cases and only in 16% of controlled cases was an 'inordinate need for affection'. The author discusses other less significant findings and examines the psychosomatic approach to the management of tuberculosis. Using Halliday's sevenfold criteria, he finds that pulmonary tuberculosis can be considered to be a psychosomatic disease.

The questionnaire method employed and the technique of controlled case study will not easily withstand criticism by depth psychologists or scientific observers. One is left with the feeling that an essentially complex matter has been oversimplified to fit hypotheses in terms of breaks in love links and an inordinate need for affection. The finding that no case of tuberculosis was without an inordinate need for affection gives this trait a specificity comparable with the tubercle bacillus which is frankly incredible. It suggests that such an indefinable trait can be found almost anywhere if one looks hard enough for it. It would be easy to elaborate further criticisms and thereby underestimate the potential value of this major study of emotional factors in pulmonary tuberculosis. This book is not easy reading, but well worth study. Its price is rather high. J. A. HARRINGTON

OTHER PUBLICATIONS RECEIVED

Social Psychology. (2nd ed.) By S. S. Sargent and R. C. Williamson. New York: Ronald Press. 1958. Pp. 649. $6.00.

Behaviorism. By J. B. Watson. (A new issue of the unamended 1930 edition.) University of Chicago Press. London: Cambridge University Press. 1958. Pp. 308. 13*s.* 6*d.*

A History of Magic and Experimental Science. Volumes VII and VIII: The Seventeenth Century. By Lynn Thorndike. Columbia University Press. London: Oxford University Press. 1958. Vol. vii, pp. 695. 80*s.* Vol. viii, pp. 808, 80*s.*

Patterns in Criminal Homicide. By M. E. Wolfgang. Pennsylvania University Press. London: Oxford University Press. 1958. Pp. 413. 40*s.*

The Introvert: a Psychiatric Study of Social Adjustment. By A. Meares. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 147. 34*s.*

Anxiety in Christian Experience. By W. E. Oates. London: George Allen and Unwin. 1958. Pp. 156. 15*s.*

Crime and Insanity. Ed. by R. W. Nice. New York: Philosophical Library. 1958. Pp. 280. $6.00.

Art Therapy in a Children's Community. By Edith Kramer. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 238. 51*s.*

Fads and Foibles in Modern Sociology. By P. A. Sorokin. London: Mayflower Publishing Co. 1958. Pp. 357. 50*s.*

Die Psychotherapie in der Gegenwart. Ed. by Erich Stern. Zürich: Rascher. 1958. Pp. 474. S.Fr. 33.

A Study in the Learning Process in the School Class Environment. By V. Heikkinen. Helsinki: Finnish Academy of Science and Letters. 1957. Pp. 54. F.mk. 260.

Comments on the Development of Teachers-in-Training. By V. Heikkinen. Helsinki: Finnish Academy of Science and Letters. 1957. Pp. 73. F.mk. 360.

The Human Situation. By W. Macneile Dixon. Harmondsworth: Penguin. 1958. Pp. 443. 5*s.*

The Theory of Evolution. By John Maynard Smith. Harmondsworth: Penguin. 1958. Pp. 320. 3*s.* 6*d.*

A Century of Darwin. Ed. by S. A. Barnett. London: Heinemann. 1958. Pp. 376. 30*s.*

Les Aveugles et la Société. By Pierre Henri. Paris: Presses Universitaires de France. 1958. Pp. 465. F.fr. 1600.

The Psychology of Medical Practice. By M. H. Hollender. London: Saunders. 1958. Pp. 276. 45*s.* 6*d.*

Psychotherapy by Reciprocal Inhibition. By Joseph Wolpe. Stanford University Press. London: Oxford University Press. 1958. Pp. 239. 40*s.*

Minnesota Studies in the Philosophy of Science. Vol. II. Concepts, Theories, and the Mind–Body Problem. Ed. by H. Feigl, M. Scriven and G. Maxwell. Minnesota University Press. London: Oxford University Press. 1958. Pp. 553. 55*s.*

Objective Approaches to Treatment in Psychiatry. By Leo Alexander. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 139. 34*s.*

VOLUME 50 MAY 1959 PART 2

AN OBJECTIVE INVESTIGATION OF THE DYNAMICS OF SEMANTIC SYSTEMS

By A. R. LURIA and O. S. VINOGRADOVA

Academy of Pedagogical Sciences of R.S.F.S.R. and University of Moscow

1

Human consciousness is a system of very intricate semantic connexions, relating to different stages of past experience and constructed on planes of different complexity. At any given moment some of these semantic systems are dominant, while others are pushed into the background and continue to exist as potentials. The objective investigation of this real relationship between actual and potential links of semantic systems, dependent on a number of conditions, continues to be a most important problem, and is as yet far from a satisfactory solution.

This problem is not solved by the investigations of conceptual thinking which are widely practised in psychology and which make use of the methods put forward in their time by Vigotsky (1956) and Goldstein (1948); these methods give a good description of semantic structures available to the subject, but do not allow one to proceed to the analysis of that real correlation between dominant and potential links of these systems, which often remain insufficiently realized by the subject himself, and which do not permit a sufficiently accurate tracing of their dynamics. Nor is this problem solved by the investigation of affective complexes which constitutes the fundamental element of psychoanalysis; the absence of an objective method remains to this day the weak point of psychoanalysis in which the description of the dynamics of meaningful associations is often transformed into a system of unfounded surmises thereby depriving it of the character of a genuinely scientific study. The proposal to make use of motor reactions, in the capacity of objective symptoms of affective complexes, which was put forward by one of us many years ago (Luria, 1932), permitted a considerable step to be taken in this field, but did not yet make it possible to proceed to that objective study of the dynamics of semantic systems which, as before, presents one of the fundamental problems of psychology.

In order to approach the solution of this problem we must consider the use of a number of objective methods, which have been formed from the study of the higher nervous activity of man and which make use of the recording of the involuntarily arising symptoms of dominant semantic systems.

One such method was applied by Schwarz (1948, 1949). Recording the change in light-sensitivity of a dark-adapted eye, resulting from a sudden illumination, she then combined this sudden illumination with the pronouncing by the experimenter of the word 'doktor' (physician); after a few combinations, this word, which up till then had been neutral, began to provoke a conditioned-reflex lowering of the thresholds of light-sensitivity. But, as experiment showed, such a lowering of the thresholds of light-sensitivity *was objectively provoked by the use of a number of other equivalent words too*, and a normal adult subject knew such words as 'vratch' and 'lekar' (both meaning physician) which sound completely different, but which are equivalent to the key word in meaning. Tests carried out to check this showed that words similar in sound but pertaining to a different sense group (e.g. the word 'diktor' (announcer) did not act as equivalent stimuli, and did not provoke a corresponding change in the thresholds of light-sensitivity). However, if the subject was given the appropriate dose of chloral hydrate, and the cortex was put into an inhibitory state—the system of equivalent links changed, and now, the patient being in a drowsy state, the word 'diktor', which is similar in sound (to the key word) began to provoke the same conditioned reflex change in sensitivity, while the synonymous words 'lekar' and 'vrach' lost their character of equivalent stimuli. Analogous tests have been carried out, using different objective indicators (the recording of salivary, galvanic skin and vascular reactions, depression of electrical activity of the brain and of the blood coagulation reactions) by a number of other investigators (Riess, 1940; Krasnogorsky, 1954; Volkova, 1953; Lacey & Smith, 1954; Marushevsky, 1957; Razran, 1949*a*, *b*; Markosyan, 1957; and others).

---

* Paper read at a meeting of the British Psychological Society, 19 August 1957.

The facts obtained in these investigations open up new ways for the objective study of the dynamics of semantic systems. They show, first and foremost, that a single verbal stimulus (i.e. the word) provokes in the subjects, not an isolated, clearly delimited reaction, but a definite *system* of connexions, and indicate that this system of links is of a strictly selective nature. They further show that, in a normal adult subject, these dominant links are meaningful links: owing to this, words included in past experience in a system and having the character of synonyms, begin 'on the spot', to provoke the same reaction, while words linked on an outward sound basis are braked and practically inactive. These facts further show that when the cortex is in a state of inhibition, this real correlation of dominant and suppressed links can alter, so that the relatively simpler sound-connexions may become dominant, and the more complex semantic connexions inhibited. Finally, and this is especially important, all the experiments show that *an objective line of investigation of the system of semantic connexions and their dynamics does exist*, and that this method is connected with the alteration in the conditions of the brain's functioning. This shows that the investigator has at his disposal the real possibility of concluding which are the dominant links of the system and how the correlation of these links actually changes with the alteration of the normal conditions.

2

The method which has been used in this study is based on data obtained in the laboratory of E. N. Sokolov (1958) and relating to the study of the components of *orienting reflexes* (Vinogradova & Sokolov, 1957).

As these observations showed, the introduction of any new stimulus (visual, auditory, tactile, etc.) provokes a distinct vascular reaction consisting of *a contraction of the finger blood-vessels and a dilatation of the blood-vessels of the surface of the head*. These vascular reactions are gradually extinguished with protracted repetition of one and the same stimulus, but recur when the stimulus is changed. They intensify when the stimulus is intensified, weaken when the stimulus is lessened, and appear especially distinctly as soon as the weak stimulus (e.g. the sound) approaches threshold intensity, becomes hardly distinguishable and consequently begins to provoke extra active orienting reaction in the subject.

All these indications permit us to evaluate the given changes as a *vascular component of the orienting reflex*. A rapid extinction of the vascular component of the orienting reflex is observed in the case of the stimulus which is irrelevant for the subject. If, with the help of verbal instruction (for instance by suggesting that the subject press a button in answer to every signal) or with the help of unconditioned reinforcement (a pain or thermic agent) the stimulus is given a signalling meaning, then the tentative reaction to it becomes considerably more stable. It is maintained during 25–40 repetitions of the stimulus, and the extinction occurs when the reaction becomes automatic.

The facts described permit the use of the method of recording the vascular components of the orienting reflex, in the question interesting us in this article—namely the objective study of the structure and dynamics of semantic systems. This was carried out by one of us (Vinogradova) in this study.

The subject, who was sitting in an armchair in a sound-proof room, was given a series of words; each word was pronounced at intervals of 20 to 180 sec. A glass tube was fixed over a finger on the subject's left hand, fitting tightly to the skin of the finger: pneumatic transmission connected it to the recording capsule on the photoplethysmograph. The change in size of the finger blood-vessels was recorded on the film. Similarly, the change in size of the blood-vessels of the head (region art. temporalii superficialis and art. frontalis) was also recorded. A capsule tightly fixed to the temporal region of the skull and attached to a special holder served as receiver.

The first part of our investigation was carried out on ten normal school-children, aged 11–15 years, and on fifteen mentally retarded children of 13–17, attending a special school. The words used in the experiment were spoken aloud by the experimenter as monotonously as possible and at one volume-level. 'Neutral' words were chosen for the experiment, i.e. words which could not provoke any special emotional reaction whatever in the subject.

As observations show, the use of different words provoked, in the normal subjects, at first orienting reactions manifested by the contraction of the finger blood vessels and by the dilatation of the blood-vessels of the head surface,* but after 15–20 different words had been given, this reaction was extinguished and the vascular curve remained on the same level when the words were given. This fade-out of orienting reflexes to any word marked the first stage of the experiment.

Against this background of extinct orienting reactions, a special signalling meaning was given to one word (in the experiments described this word was 'koshka' (cat)). The subject was asked to press a button with his right hand every time he heard this word; on the photographic record the motor reaction of the right hand was registered and also the vascular reaction of the left finger.†

The experiment showed that, in these conditions, *every time the signal word, 'koshka', was given it evoked in a normal subject not only the motor reaction of the hand, but also a steady orienting vascular reaction,* which was maintained for a long time (during 20 to 40 repetitions), whereas the giving of other 'neutral' words provoked neither a motor nor a vascular reaction. Such a distinct isolation of one signal word, evoking a stable vascular reaction, allowed us to begin the main experiment.

The first question before us was the following: is the orienting vascular reaction evoked

---

[*] For simplicity of exposition in this part of our report we shall use only one of these indicators—the contraction of the finger blood-vessels. We shall specially examine the relationship of both indicators later.

[†] Naturally the first thing that was established was that the pressing of the button with the right hand evoked only a momentary occurrence of vascular reaction in the finger of the left hand and had no influence on the vessel reactions in further experiments.

only by the *given* signalling word, or does the same effect extend to a *whole system of words*, and can it be evoked by a number of other verbal stimuli which—*without evoking an open motor reaction*—could however *evoke a vascular reaction*? In the event of such a thing taking place, we would be able to work out objectively *a system of semantic connexions characteristic of a given subject in the circumstances under investigation*.

Observations drawn from corresponding experiments make it possible to answer this question.

If, under the conditions just described, we gave the subject an absolutely neutral word (having nothing in common, either with regard to sense or to sound, with the key word 'koshka' (cat); e.g. 'steklo' (glass), 'karandash' (pencil), 'oblako' (cloud), etc.) we did not get a vascular reaction. If we gave the subject a word connected in meaning to the key word ('kotyonok' (kitten), 'mysh' (mouse), 'zhivotnoye' (animal), 'sobaka' (dog)) *though there was a complete absence of motor reaction, there occurred a distinct contraction of the finger blood-vessels, which pointed to the presence of an orienting reaction*. This gives us grounds for assuming that the given words were not without significance, for the subject. A heightened orienting reaction to these words can be explained only by the existence of a functional semantic connexion with the basic signal word, formed in the past experience of the subject. This fact was constant, and words having a direct connexion with the signal word provoked a relatively more manifest vascular reaction, whereas words of a general nature, and linked less directly and obviously with the critical word, provoked a less evident vascular reaction. In contrast to this—words not having a sense link with the critical word and bearing only a superficial phonetic resemblance to this word (e.g. 'okoshko' (window), 'kroshka' (crumb)) did not provoke a vascular

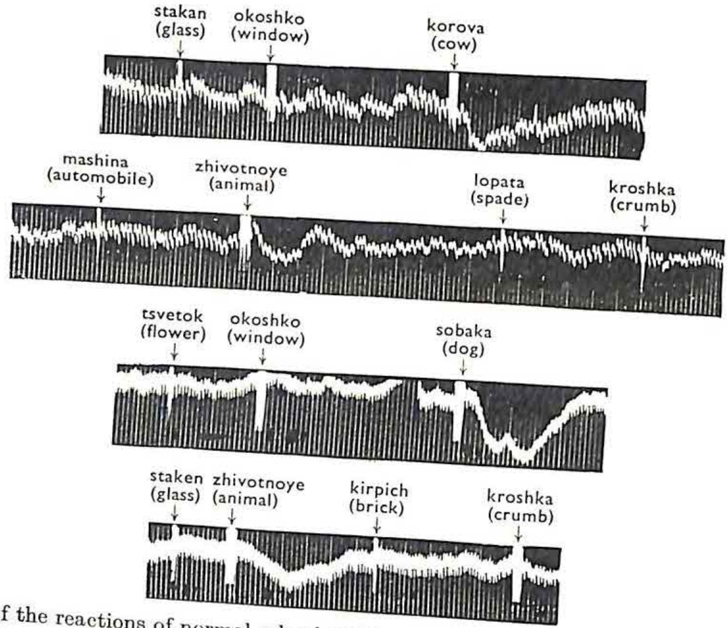


Fig. 1. The nature of the reactions of normal school-children of 11–12 years to word stimuli. Words linked in meaning with the signal word 'koshka' (cat), 'korova' (cow), 'sobaka' (dog), 'zhivotnoye' (animal) evoke reactions; words similar in sound, 'kroshka' (crumb), 'okoshko' (window), have no effect. In the figure the plethysmogram of the finger is recorded. The moment of application of word stimulus is registered by vertical lines. The thin vertical lines are the recording of time in seconds.

reaction, thereby indicating clearly that, in a normal subject, they are not included in his actual system of sense links. Fig. 1 gives examples of the facts described.

These experiments allow us to draw important conclusions. They show that, in evoking an orienting reaction to one word, we bring about, at the same time, *a whole system of meanings*. This can be revealed objectively in those cases where the subject does not display any outwardly manifest voluntary reactions to words pertaining to this system. These experiments further corroborate the fact established earlier, namely that the semantic is the dominant system, while sound resemblances, pertaining to the outward sound aspect of the word, are not included in the dominant system. Finally, the experiments allow us to take a few steps towards the objective analysis of the stability of these connexions, distinguishing the group of words which reveal a closer link with the given key word, from those revealing a more remote link. In this way, we are brought nearer to an objective analysis of the structure of this system of semantic connexions.

On the basis of the experiments carried out we can consequently conclude that the whole semantic system obtained in our 'model' experiment acquires the following structure: in the centre of this structure stands the *key word*, which has a signalling meaning and evokes both a motor and an orienting vascular reaction. This 'nucleus' is surrounded by a whole system of *potential sense-links*, to which pertain words which do not evoke an overt motor response, but which do evoke only an orienting reaction. Finally, outside of this system lie those neutral words which do not evoke any motor or vascular reaction at all. In the normal subject, words not having a semantic connexion with the critical word, but resembling it in sound are also related to these 'extra-systemic' signals.

3

However, the question arises: how is the described semantic system changed when certain disfunctions of the brain are present? To this end experiments were carried out on mentally retarded children suffering from oligophrenia,* in different degrees (debility and imbecility) and being taught in special schools. As previous investigations of the higher nervous activity of these children have shown (Luria, 1956, 1958), under-development of the higher functions of the speech system is characteristic of such children.

Experiments in this series gave results which differed sharply from the above-described experiments with normal school-children.

The system of verbal connexions which we were able to ascertain from our subjects was, as a rule, also of a selective nature. But the nature of this selectiveness was different and less clearly manifest than in normal persons, and the more profound disturbance of the mental activity and the corresponding disturbance of the intellectual process, the greater was the disturbance of the system of links elicited in these cases.

In children having a relatively less manifest form of oligophrenia (oligophrenia in the debility stage—and mainly the older age-group, 15–17 years old) we could, as a rule, observe a peculiar fact: the system of dominant connexions comprised *both words near to the key word in meaning and words similar to it in sound*. Vascular orienting reactions were evoked here, both by the word 'sobaka' (dog) which is linked in meaning with the key signal word, and by the words 'kroshka' (crumb) and 'okoshko' (window) which have merely a superficial phonetic resemblance to it. However, a motor reaction resulting from verbal instruction was evoked only in response to the signal word, 'koshka'. This fact has the characteristic feature that, if vascular reaction to the latter words (similar in sound) occurred after a short latent period, being thus immediate in character, then the vascular reaction to words having a sense-link with the key word occurred after a significant latent period (up to 7 or 8 sec.), and hence arose as a result of another considerably less direct process. This fact illustrated in Fig. 2 gives grounds for important conclusions with regard to both the smaller selectiveness of the system of word links in the mentally retarded children and to the different actualization of sound and meaning systems which are characteristic of a pathologically altered cerebral activity.

---

* As regards oligophrenics, we took only those children having undergone, either while still in the womb or during early development, a severe cerebral lesion as a result of which they showed mental defect.

We are able to observe even more significant changes in the system of connexions, in experiments with children, having the defect in a more serious degree—or with debile children of the younger age-group (12–14 years).

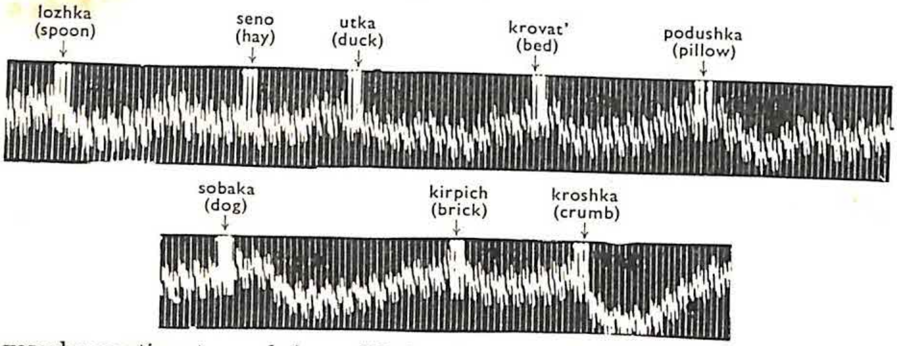


Fig. 2. The vascular reactions to words in a mild oligophrenic (a debile of 16 years). Vascular reactions occur both in semantic and in sound connexions.

As a rule, words close to the key word in meaning, 'kryssa' (rat), 'sobaka' (dog), 'zhivotnoye' (animal), did not provoke orienting vascular reactions and consequently did not generally come into the sense-system, worked out by us (only those words having a direct connexion to the key word proved the exception; such as, for example, the word 'kotyonok' (kitten). Instead, words having a sound similarity to the critical word (e.g. 'kroshka' (crumb), 'kryshka' (lid), 'kruzhka' (mug), 'okoshko' (window), etc.) continued to evoke a vascular reaction; i.e. the nature of the links discovered was directly opposite to that observed in the norm. It is interesting that the word 'myshka' (little mouse) also evoked the active tentative vascular reaction; but when only the root of this word was left ('mysh' = mouse) and the sound resemblance ('--shka') to the key word eliminated—it ceased to evoke the vascular reaction.

In even more serious forms of oligophrenia (imbecility) the specificity of the responsive motor reaction disappears. Despite precise instructions to press the button only in answer to the word ('koshka') the subjects reacted in this way to words similar in sound to the key word. In this way pathological deformation of the verbal system is associated with disturbances of the voluntary motor reaction, which reveals the same defects as the involuntary vascular component of the orienting reaction. But even in these cases the motor reaction remains less diffuse in character than the vascular. Thus if a vascular reaction is evoked even by words so dissimilar to the key signal, such as 'chashka' (cup) or 'kruzhka' (mug), then the motor reaction occurs only with those words most similar in sound (e.g. 'kroshka', 'okoshko').

All these facts, an illustration of which is given in Fig. 3, clearly indicate that *the system of links resulting from a word is distinctly altered and pathological in character, and that the dominant position of selective semantic connexions disappears here*, while *sound*-resemblance, not of a selective sense-link nature, begins to occupy the leading position. It is characteristic that this disturbance of the semantic system is evident here too, and the system ceases to be clearly divided into the group of *central signals*, provoking a motor reaction, and the group of *peripheral verbal signals* which do not provoke a motor reaction, but which only result in the evocation of distinct vascular reactions.

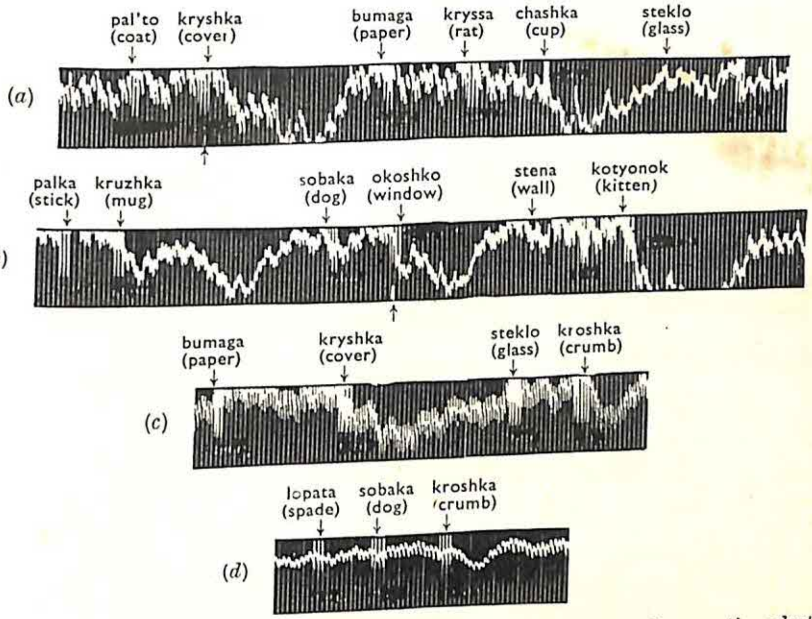


Fig. 3. (*a*) The reaction to words of an imbecile (16 years). Not only vascular reactions but also arbitrary motor reactions are disturbed. Explanation given in the text. (The motor reaction is registered by a vertical line ↑ in the lower part of the tracing.) (*b*) and (*c*) are examples of vascular reactions in child oligophrenics (debiles of 12–14 years). While there is an absence of reaction to words having semantic connexions with the signal word, reactions to words similar in sound to the signal are observed.

4

An objective analysis of the semantic system is not limited to the study of the structure which we have just described. It permits us to go on to the analysis of the *dynamics* of these connexions and the changes in them which depend on the state of the subject and the conditions of the experiment.

In the dynamics of the experiments carried out on our subjects, who revealed the described pathological change in the system of meaning connexions, we are able to record two factors which act in a mutually opposing direction. One of these is the factor of *training* or *exercise*, which intensifies the selective nature of the connexions; the other is the factor of *fatigue* or *exhaustion*, which leads to an inhibitory state of the cortex and disturbs the system of selective connexions.

Among our subjects were a number of slight oligophrenics, who in the original experiments revealed characteristic pathological alterations in the system of verbal connexions. There was an absence of vascular reaction to stimuli similar in meaning, in the presence of vascular reactions to words similar in sound. But in the course of the experiments, with repeated reinforcement of the signal word, 'koshka', there was a material change in the situation. The subject began to form a system of selective semantic connexions, evident in that words having an outward sound resemblance to the key word ceased to evoke a vascular reaction, while words linked in meaning with the key word acquired a signalling character. An example of such an effect of training is shown in Fig. 4.

Fatigue, effecting an inhibitory state of the cortex acted in the opposite direction. This could be discovered by carrying out experiments twice in the course of a day—morning and evening. Thus in the experiment with an oligophrenic of the older age-group (18 years old) carried out in the morning before he began his school work, normal relations in his word system were ascertained. Then an experiment of the same sort was repeated after 6 hr. of work in class. In this experiment were revealed a disintegration of sense-links and the occurrence of reactions to words resembling the key word in sound; i.e. a more primitive system of connexions (Fig. 5) was shown. The same phenomenon was observed in the presence of headaches, influenza, and other disturbances of the general condition of the subject.

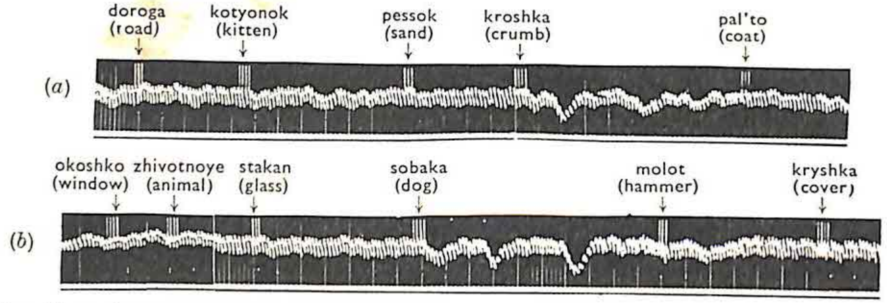


Fig. 4. The effect of training on the dynamics of semantic system in an oligophrenic (a debile of 12 years). Sound-similarity gives place to adequate semantic connexions. (*a*) First experiment; (*b*) fourth experiment.

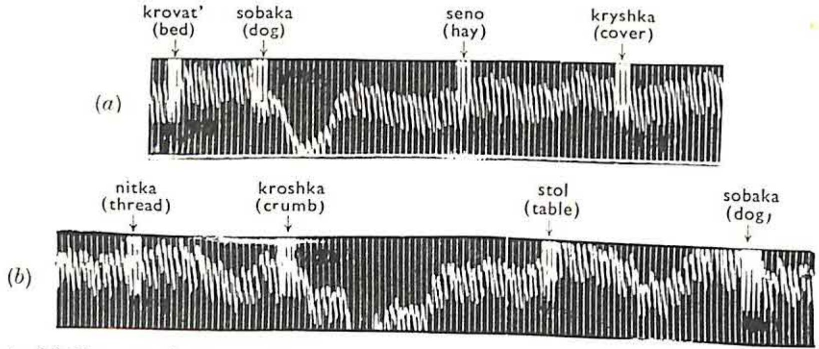


Fig. 5. Effect of fatigue on the semantic system of an oligophrenic (a 17-year-old debile). Adequate semantic system gives way to primitive sound connexions. (*a*) Before lessons; (*b*) after lessons.

The dynamics of semantic systems discovered in our experiment is not defined by the factors of the general condition of the subject alone; they may depend on a *number of inner conditions of experiment*, on the basis of which we may have the possibility of *controlling the dynamics of the systems investigated*. The *inclusion of the word in a definite context* can serve as one of the clearest examples of such a control. We have already said that a fact, characteristic of deeply affected oligophrenics, is the appearance of primitive verbal connexions owing to an outward sound resemblance as a result of which the word 'okoshko' (window) having merely an outward sound resemblance to the signal word, 'koshka' (cat), begins to evoke vascular reactions. But if, by the introduction of this word into a definite, specially selected series of words, we intensify its own semantic connexions, the sound resemblance to the signal word can be inhibited and this word stops evoking vascular reactions. We give an example of such control of the sense-links in Fig. 6. Without interrupting the experiment with the oligophrenic subject in whom the word 'okoshko' provoked a tentative reaction by virtue of its resemblance in sound to the word 'koshka', we introduce this word into a context of words relating to other parts of the house—giving consecutively the words 'dom' (house), 'dver' (door), 'stena' (wall), 'okoshko' (window). Under such circumstances, the proper semantic field of this word is intensified, its sound resemblance to the signal word loses its significance, and the same word now ceases to evoke an orienting reaction linked with the signal word.

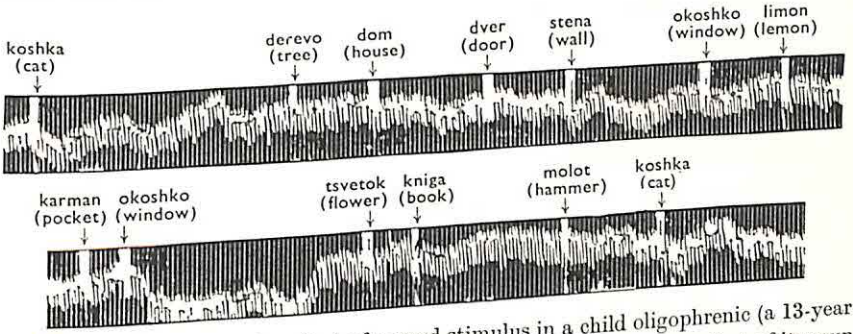


Fig. 6. The influence of 'context' on the effect of a word stimulus in a child oligophrenic (a 13-year-old debile). The introduction of words into the corresponding semantic system brakes the effect of its sound similarity with the signal word (explanation in the text).

The given experiment, being an attempt at active influence on the structure of the word-system which has been revealed, and at controlling the dynamics of this system, may suggest new lines for the investigation of different aspects of the conscious organization of sense-activity in the normal as in abnormal subjects.

5

So far we have only considered the general indicators of the presence of semantic systems which we were enabled to approach through the method of recording of the vascular components of orienting reactions.

Can we not, however, take a further step and try to introduce a further qualitative analysis into the structure and dynamics of those potential links which we have just described?

For this purpose we must modify the method and use such means as would allow us to *differentiate variants* of such potential sense-links, not only in quantitative terms (more or less intensive reaction) but when possible in *qualitative* terms also. The method which one of us (Vinogradova, in collaboration with Heissler*) had used was found suitable for these problems too.

In the variant of the method used for these purposes, we discarded the motor reaction of the subject, but replaced orienting reinforcement (conditioned by giving special significance to one of the verbal stimuli) with a *pain* reinforcement. The given method is based on data obtained earlier in our laboratory by Vinogradova & Sokolov (1957). As these experiments showed, thermic stimuli, which provoke specifical thermoregulatory vascular reactions, and also pain stimuli evoked characteristic changes in the blood vessels of the hand and of the surface of the head. These changes are expressed in the parallel contraction of the blood-vessels of the head and of the hand under the action of pain and cold and in the dilatation of the said blood-vessels under the action of warmth. However, these specific reactions do not arise at once, and in the first experiments they are preceded by non-specific (alike in all cases) *orienting* reactions consisting of the *dilatation of the blood-vessels of the head and the contraction of the blood-vessels of the finger*. Only thereafter does this non-specific reaction fade out and change into the specific (adaptational or defensive) reaction mentioned above, during which the reaction bears the same character both in the blood-vessels of the head and in the blood-vessels of the hand (see Fig. 7).

---

* A report of this work is in preparation.

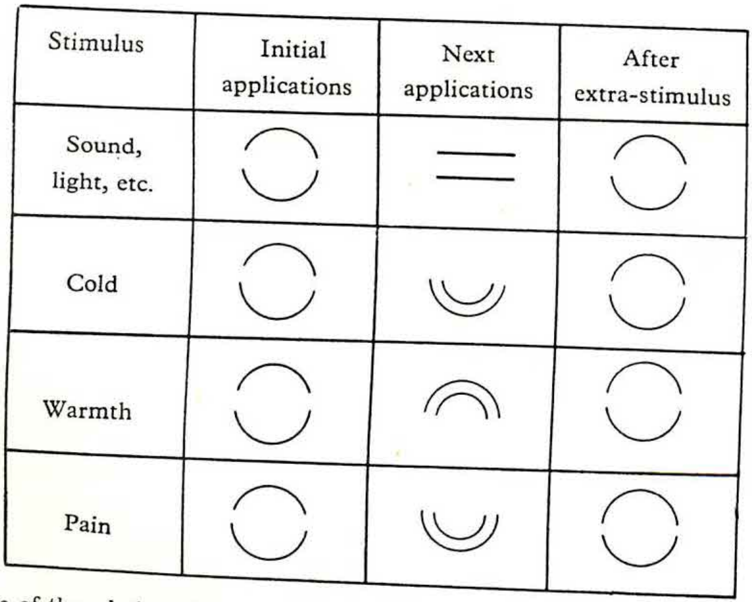


Fig. 7. The scheme of the relation of non-specific orienting and special reactions to different stimuli during recording of the plethysmogram of the blood-vessels of the finger and the surface of the head.

This basic fact was used by us for the further objective differentiation of the dynamics of semantic system. This series of experiments was carried out on seven normal adult subjects. When the preliminary orienting vascular reactions had faded out on all words, we began to accompany one of them (in this series of experiments it was the word 'skripka' (violin)) with a painful electric stimulation applied to the skin—a current of 45 volts. In accordance with our earlier data the reinforcement of the word 'skripka' (violin) with a current at first evoked a reaction of an orienting nature (contraction of the finger blood-vessels and dilatation of the blood-vessels of the head) which was gradually replaced by the specific pain reaction (parallel contraction of blood-vessels of the head and hands) only after 18–25 repetitions of such combinations.

At this phase we went over to the basic tests. Among various 'neutral' words we began to introduce words having semantic connexions with the basic stimulus ('skripka') or having a sound similarity with it.

It became clear that the words we used provoked reactions of varied character. A fairly wide group of words, in spite of the fact that they had not been reinforced by current, *evoked exactly the same reaction as the basic stimulus, that is the specific pain reaction of parallel contraction of the blood-vessels of the hand and the head.* As the analysis shows, this group of words obviously had a more direct close link with the key word 'skripka' (violin). It included such words as 'skripach' (violinist), 'smychok' (bow), 'struna' (string), 'mandolina' (mandoline), 'kontrabas' (double-bass) and several other names of stringed instruments. Besides this group of words there was another still wider group which also evoked vascular reactions, *although these reactions were not defensive but orienting in character, that is were expressed in the dilatation of the blood-vessels of the surface of the head and the contraction of the finger blood-vessels.* To this group belonged: (*a*) names of stringless musical instruments ('akkordeon' (accordion), 'klarnet' (clarinet), 'baraban' (drum), etc.); (*b*) other words connected with music ('sonata', 'kontsert' (concert), etc.); (*c*) in a number of cases the same reaction was evoked by words close to the key-word in sound ('skrepka' (paper-clip)). In addition to this, 'neutral' words did not evoke any changes in the vascular curve (e.g. 'shkaf' (cupboard), 'sapog' (boot), 'oblako' (cloud), etc.).

In the further conduction of the experiments, this phase of *primary systemic generalization* changed over into a phase of *concentration*: only the key word ('skripka') reinforced by the current continued to provoke a distinct specific (pain) reaction. *All* words close to it in meaning began to evoke an orienting vascular reaction. The group of words, which in previous experiments had evoked an orienting reaction (including also words without a semantic connexion with the key word, although similar in sound to it) ceased to evoke any vascular reaction whatsoever and became completely neutral. Finally, in further experiments came the stage of *full concentration*—only the key word as before kept its defensive action; all the remaining words, including even the closest in meaning, went over to the category of neutral words. An example of such dynamics of vascular reactions to various words is given in Fig. 8.

It is characteristic that the facts described above, i.e. the dependence of the objective vascular reactions to a given word on its context and the possibility of directing the signal meaning of words, are repeated in a new form here also. The word 'truba' (pipe or chimney) which had given a distinct orienting reaction, when said out of context, ceases to give it if the word is included in the context 'dom' (house) 'pechka' (stove), 'krysha' (roof), taking on by that very fact the meaning 'chimney' (Fig. 9).

All these facts allow us to come to two conclusions.

The objective study of vascular reactions under the conditions of our experiments allowed us to affirm the fact that all the words are divided into three groups: (1) *the nucleus of the semantic complex*, to which is related the key word and words in direct semantic proximity to it (they evoke a specific pain reaction); (2) *the periphery of the semantic system* to which are related words linked less directly with the key word (evoking a non-specific orienting reaction); and (3) *neutral words*, which in our experiment did not evoke any specific or orienting reactions.

The correlation of these groups at various stages of the experiments may be different, and if at the beginning of the experiment the nucleus of the semantic system has a relatively generalized character, then later it becomes concentrated and only the key word continues to provoke a specific reaction, while the remaining words, which earlier were included in this nucleus, move into the semantic periphery and begin to evoke only a non-specific orienting reaction.

In the method used we gain consequently the possibility of taking one more step along the path towards the definition of the *structure and dynamics of semantic system*, to the objective study of which we have dedicated our research.

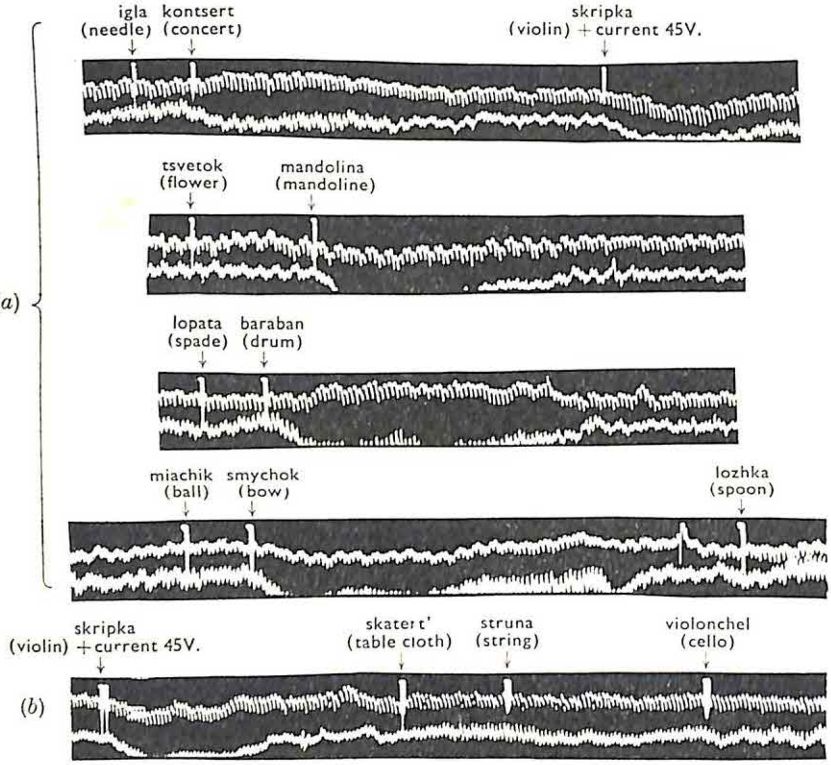


Fig. 8. (*a*) The initial differentiation of word stimuli during the working out of the conditioned defensive reflex to the word 'skripka'. The words 'smychok' (bow), 'mandolina' (mandoline), evoke a defensive reaction; the words 'kontsert' (concert), 'baraban' (drum)—a reaction of the orienting type. (*b*) The stage of final concentration (fifth experiment)—only reaction to the reinforced word remains. In each record the upper plethysmogram is that of the blood-vessels of the surface of the head, the lower plethysmogram is that of the finger blood-vessels. The double horizontal line register sthe unconditioned reinforcement by electric shock.

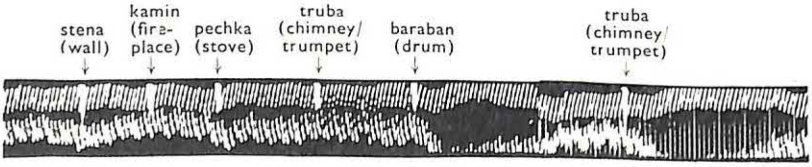


Fig. 9. The influence of a semantic text on the appearance of different connexions of one and the same word, from the phonetic point of view (homonyms). Explanation in the text.

6

We have traced the natural dynamics of semantic systems obtained by us in those conditions where we reinforced only one key word. We saw that by the repetition of the experiments in accordance with a definite plan, we reach a certain stage when the extended connexions of the given word ceased to appear, since the limit of concentration of the system has been reached. However, at any moment we can restore the whole semantic system being investigated, in all the fullness of its connexions. For this it is sufficient to reinforce with the current one or two words, besides the key word 'skripka', *relating to the same system* ('mandolina', 'gitara' (guitar), etc.). A single reinforcement of these words with the current leads to the immediate *restoration of the whole active system of words*, linked by direct and indirect connexions with the basic stimulus. When this is done there appears again the field of words closer to the basic core, which evoke defensive reactions, a semantic periphery, evoking only orienting reactions and a neutral zone which does not change its indifferent character as regards vascular reactions. Fig. 10 gives an example of such a change.

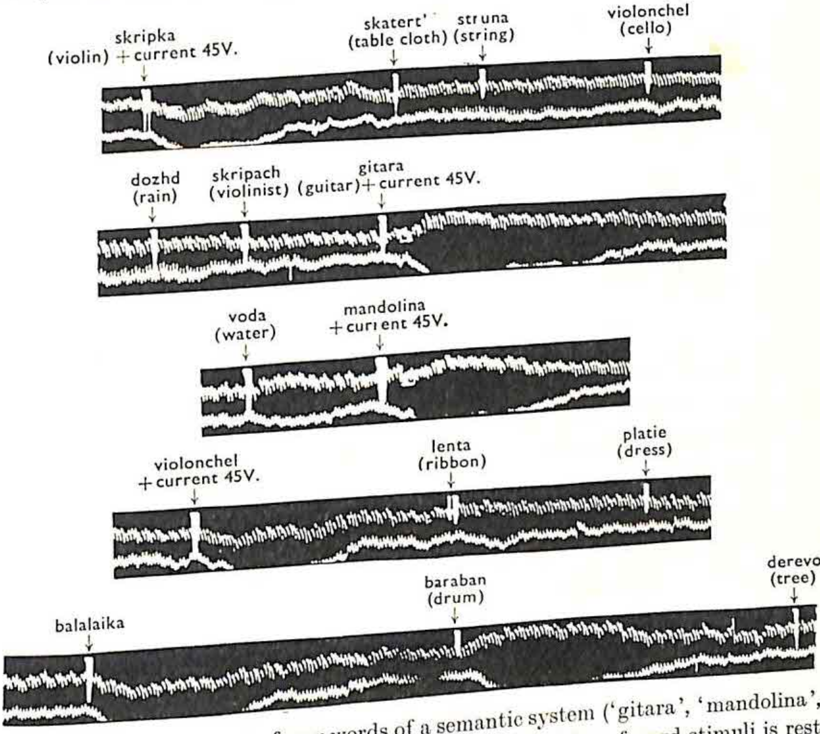


Fig. 10. The effect of the reinforcement of new words of a semantic system ('gitara', 'mandolina', 'violonchel') on the participation of a reaction. The initial differentiated effect of word stimuli is restored.

The systemic nature of the restoration of differentiated vascular reactions becomes even more obvious if, in the capacity of a control, we introduce the following experiment. At the stage when only the basic word 'skripka' (violin) preserves its action, *we reinforce with the current another word, having no relationship with the system investigated earlier, for example the word 'korova' (cow).* Such a method does not have a restorative influence on the earlier system (the 'skripka' system)—its components remain indifferent as regards vascular reaction. The new reinforced stimulus creates *its own system* in which, as we already know, the character of the reactions is differentiated according to the degree of proximity to the basic irritant (such as the words 'byk' (bull), 'telyonok' (calf), etc., evoke here a defensive reaction, the words 'stado' (herd), 'moloko' (milk), 'sobaka' (dog) evoke an orienting reaction, and so on). Thus these two systems, foreign to each other as regards their semantic fields, turn out to be, as it were, absolutely isolated in their objective manifestations. However, this is not entirely correct, and with the help of a special variation of our technique we can discover the mutual influence of the two semantic groups created by us.

If we begin to reinforce the word 'korova' (cow) before the final differentiation of the 'skripka' system, that is when the elements nearest it continue to evoke orienting reactions, we will see that the mutual influence of the two systems does, in fact, take place. If after several reinforcements of the word 'korova' we use the words of the previous system, which had just evoked distinct orienting reactions, they turn out to be inactive.

Thus, the rise of a new dominant semantic system sharply differing from the system which had existed earlier leads to a *temporary inhibition of that system*, to the breaking of its own peripheral links, though the core (defensive reaction to the word 'skripka') is relatively well preserved.

Alternating periods of the reinforcement of the word 'korova' with periods of reinforce-

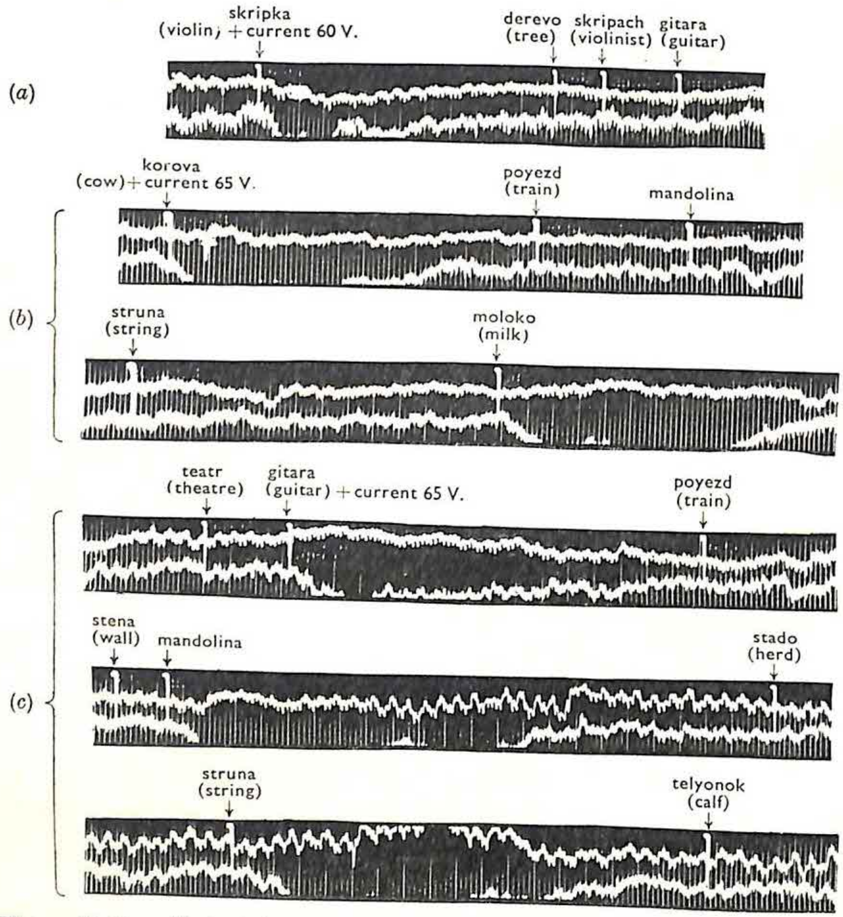


Fig. 11. The conflicting effect of the reinforcement of another semantic system ('korova'—cow). With the
appearance of a new group of signal stimuli ('moloko'—milk, 'telyonok'—calf) the action of the stimuli
of the initial sense system ('skripka', 'gitara', 'mandolina', 'struna'—string) is braked. A new reinforce-
ment of the initial semantic system (gitara + the electric shock) leads to the suppression of reactions to the
words of the new complex ('stado'—flock, herd, 'telyonok'—calf) and the renewal of the former complex
of links (mandolina, struna). Simultaneously, a conflicting state arises manifesting itself in the change in
character of the vascular curve. (*a*) Reinforcement of the word 'violin'. (*b*) Reinforcement of the word
'cow' (twice). (*c*) Reinforcement of the word 'guitar'.

ment of the word 'skripka', we can repeatedly change the nature of the action of the test words, which relate to the one and to the other system. The activation of the system of the word 'skripka' (by means of using this word, reinforced by the current, two or three times in succession) leads to the suppression of the links of the 'korova' system, and vice versa.

The given semantic systems turn out to be mutually exclusive, the actualization of one sense complex leads to the repression of the other.

This work on the repeated reconstruction of the system of sense-links does not pass by without leaving its mark on the subject. Respiratory waves, appearing on the plethysmographic curve, general volume oscillations and irregularity of the pulse point to the difficulty with which this reconstruction arises and to that irradiated disorganization of the vegetative processes, which is a symptom of the difficulties which occur. Fig. 11 gives an example illustrating this state.

Experiments with the reversal of semantic fields and with the study of the objective symptoms of this reversal open up important perspectives of research. They permit the objective approach to the dynamics of the conflict of two systems, of the 'suppression' of a system, which results from such a conflict and of that general reaction to the switching of dominant systems, which were studied by Pavlov in animals, and in the early works of one of us (Luria, 1932) were just tentatively traced in man.

7

We have examined the facts which were obtained with the help of the objective method of studying vascular reactions to the appearance of a signal word.

We have still to discuss a final, but very important, question—what is the relationship of the semantic system appearing in the objective symptoms described above, to the system of conscious connexion which may figure in the *verbal account of the subject*? The fact that in the first series described above vascular reactions also accompanied those signals which by themselves did not evoke motor reactions compels us to think that *by no means all the links reflected in them occur equally consciously*, and that in many cases *vascular reactions reveal potential systems of links more fully than the conscious account of the subject is able to.*

In order to verify this basic question we interrogated our subjects at the termination of the described series of experiments.

The results obtained from such an interrogation were sufficiently uniform. The subjects, as a rule, *were unable either exactly to formulate the aims of the experiment, or to designate exactly that group of words evoking definite vascular reactions*. An extract from the report of the interrogation of one of the subjects may serve as an illustration of this position.

Subject A. S. (scientific worker, 25 years old). Experiment no. 4, 4 May 1957

Tell us what happened during the experiment?

All kinds of words were said, without any system, in my opinion. No, in general, you probably have a system, but I do not know it—the words were very different. After certain words, the shock was given. These words were 'gitara' and 'skripka'.

What other words do you recall?

'Doroga' (road), 'korova' (cow), 'lozhka' (spoon), 'salat' (lettuce), 'mandolina' (mandoline), 'akkordeon' (accordion), 'kosa' (scythe); 'koryto' (trough)—that was the first word.

What did you feel, what was your attitude to the experiment?

I took it as a matter of course, and did not react to anything. When there was a shock after new words I was surprised, but not much.

Can you not say what was the aim of this work, was there no special choice of words, or links between them?

I don't know....I don't even imagine....You see the words were so varied....

Perhaps some words occurred more often than others?

I did not pay any attention to that. In my opinion the word 'lozhka' (spoon) was somehow repeated. Well, 'skripka' was repeated, of course, and 'gitara'.

Can you not apply a single word to a section of the words used?

Well, I don't know....I believe they were musical words, the names of musical instruments, but in general they were very varied.

You are right. You were given names of musical instruments. But why were you given these particular words?

Probably...of course, you gave me a shock after the word 'skripka'...I can guess...oh, and after 'gitara' too!

Has it occurred to you, or did you guess it during the experiment?

No, I didn't think of it then; for me there was no connexion of words.

Thus the system of links revealed by us, is not of a clear conscious nature, and *it is reflected much more fully in the system of vascular reactions than in the conscious account of the subject.*

The semantic analysis of connexions formed in the subjects is confirmation of this peculiarity. It shows that the semantic system which is formed can only with difficulty be fitted into any strict logical system of concepts. Thus, for example, the group of named musical instruments, which evoked, in the first stages of the experiments the same defensive reaction as the word 'skripka' could be joined in one logical category as 'stringed instruments'. But, together with this, the word 'arfa' (harp), for example, never evoked a reaction of a similar sort, which compels us to suppose that the principle of the similarity of a certain situation or the outward shape of the instrument also played a substantial part in the process of the revelation of the semantic field under investigation. Hence it is evident that it would be a great mistake to approach the real semantic field arising in the brain cortex of the subject as a system of logical categories, which is sometimes of a broader, sometimes of a narrower character.

The absence of adequate identification of the semantic field obtained in these experiments revealed objectively in the experiment agrees completely with the analogous data of Lacey & Smith (1954), who had used as indicators of verbal connexions which occurred changes in the pulse rate and the skin-galvanic reflex. In their experiments also, the subjects did not reflect in their report connexions which it was possible to ascertain by means of objective indicators. Another point of view is held by Razran (1949*b*), who considers that the generalization of stimuli occurs as a result of the evaluation of a new stimulus on the principle of similarity-difference, and can be revealed as a result of introspection. However, he himself shows (1949*a*) that the introduction of various instructions in experiments in which the generalization of word stimuli was being studied did not reveal the considerable discrepancy of results, which forced him to come to the conclusion that the influence of voluntary and thought processes on the gradients of generalization is insignificant.

We believe that in the given experiments we have revealed semantic fields or systems formed in past experience and automatic in an adult person. These fields or systems, which are not perceived at every given moment, form the background of any thought process. With the help of instructions (e.g. in the classification experiments) we can at any moment raise these systems to the level of consciousness, but in ordinary conditions this does not occur. The given method makes it possible to approach, without the active participation of the subject himself, the complexes of consolidated semantic systems which have been formed in the process of social speech experience and which are at the same time peculiar to the individual.

REFERENCES

GOLDSTEIN, K. (1948). *Language and Language Disturbances.* New York: Grune and Stratton.

KRASNOGORSKY, N. I. (1954). *Physiology of the Development of Children's Speech. Collected Works,* vol. I, 462–81. Moscow.

LACEY, J. I. & SMITH, R. L. (1954). Conditioning and generalisation of unconscious anxiety. *Science,* **120,** 1045–52.

LURIA, A. R. (1932). *The Nature of Human Conflicts.* New York: Liveright.

LURIA, A. R. (1956, 1958). *Problems of the Higher Nervous Activity of Normal and Abnormal Children,* vol. I (1956); vol. II (1958). Moscow: Acad. Pedag. Sciences.

MARKOSYAN, A. A. (1957). On the interaction of two signal systems in involuntary processes. *Proceedings of Conference on Psychology, 1–6 July,* 1957, pp. 152–60. Moscow: Acad. Pedag. Sciences.

MARUSHEVSKY, M. (1957). On the interaction of two signal systems in vascular reactions. *Vopros. Psikh.* **1,** 78–87.

RAZRAN, G. (1949*a*). Semantic, syntactic and phonetographic generalisation of verbal conditioning. *J. Exp. Psychol.* **39,** 642–52.

RAZRAN, G. (1949*b*). Stimulus generalisation of conditioned responses. *Psychol. Bull.* **46,** 337–65.

RIESS, B. F. (1940). Semantic conditioning involving the Galvanic skin reflex. *J. Exp. Psychol.* **26,** 238–40.

SCHWARZ, L. A. (1948, 1949). Knowledge of the word and its sound form as a conditioned stimulus. *Bull. Exp. Biol. Med.* **25,** 292–4; **27,** 412–15.

SOKOLOV, E. N. (1958). The orienting reflex, its structure and mechanisms in *The Orienting Reflex and Exploratory Behaviour,* pp. 111–22. Moscow: Acad. Pedag. Sciences.

VIGOTSKY, L. S. (1956). *Selected Psychological Studies.* Moscow: Acad. Pedag. Sciences.

VINOGRADOVA, O. S. & SOKOLOV, E. N. (1957). A study of the reactions of blood vessels of the hand and head as components of certain unconditioned reflexes. *J. Physiol.* (U.S.S.R.), **42,** 57–67.

VOLKOVA, V. D. (1953). On certain characteristics of the formation of conditioned reflexes to speech stimuli in children. *J. Physiol.* (U.S.S.R.), **39,** 540–8.

(*Manuscript received* 11 *January* 1958)

THE STROBOSCOPIC PATTERNS

I. THE DARK PHASE

By J. R. SMYTHIES*

Psychological Laboratory, University of Cambridge†

A detailed analysis of the dark phase patterns of Brown and Gebhard produced by looking at a uniform flickering field has been made. The phenomenon may be classified into various types, general features of the patterns have been described and their predominant colours have been evaluated quantatively. The effect of monochromatic light was explored and quantitative experiments were conducted to determine in the optic system of which eye the patterns arise. An argument is presented that they arise in the optic system of the externally dark eye and the mechanism by which this may be brought about is discussed.

I. Introduction and outline of previous research

The fact that complex visual perceptions of an illusory nature can be evoked by looking at a rhythmically flashing uniform field has been known since the discovery of the phenomenon by Purkinje (1823), but only recently, following the work of Grey Walter, (1950) has much attention been paid to it. The most significant paper on this subject in recent years was published by Brown & Gebhard (1948). In this they presented their interesting discovery that the phenomenon is divisible into two quite different kinds, which they called the 'bright' and 'dark' phases. They used only single eye illumination in two subjects and reported that these two phases alternated in a manner very reminiscent of retinal rivalry. The 'bright' phase was characterised by geometrical patterns and the dark phase consisted of 'a writhing mass of violet flecks against a yellow-green background. Their constant motion was rapid and curvilinear.' Because of the alternation of the two phases they suggested that the dark phase patterns are 'a function of the externally dark eye and that the bright phase occurs in that portion of the optic sector served by the directly stimulated eye'. They then tested this hypothesis by illuminating the externally dark eye at intervals with a weak, diffuse illumination and they observed that 'whenever the bright phase was dominant, stimulation of the right (dark) eye brought instant dominance of the dark violet and yellow-green pattern. They constructed three hypotheses to explain how a light stimulus in one eye could evoke sensations seen with the other eye. They thought this could be done (*a*) by a central effect, (*b*) by an effect conducted to the other eye by centrifugal fibres in the optic nerve, or (*c*) by light leaking across to the other orbit via the nasal and orbital tissues in between. They did not attempt to make any choice between these various hypotheses but merely indicated that they considered the latter to be the most conservative. This present paper will be concerned with the dark phase patterns, for, if Brown and Gebhard are correct in supposing that they originate in the optic system of the externally dark eye, this becomes a phenomenon of great interest (if the hypothesis (*c*) can be excluded): I do not think that it has been previously suggested that activity in the neurones connected to one eye can induce detectable activity in the neurones connected to the other eye. Something like this must of course happen in the case of the still unknown mechanisms whereby the two uniocular fields are fused in normal stereoscopic vision. But here we are dealing with something different, for in the case of the dark phase patterns detectable activity may be induced in neurones not under stimulation by their own sense organs. This present investigation was therefore undertaken using more observers and subjecting the patterns to a closer analysis.

---

* This work was conducted during the tenure of a Nuffield Fellowship in Medicine.

† Now at Worcester Foundation, Shrewsbury, Mass., U.S.A.

II. MATERIAL AND METHODS

Two sources of intermittent light were used. (i) A beam of light from an Aldis 500 watt slide projector was interrupted by an episcotister driven by a variable-speed motor. This beam illuminated a frosted glass screen. The subject sat behind the screen with his eyes 10 cm. from it so that the illuminated region covered most of his field of vision. With the rotating sector in position the illumination of the screen was 18·7 millilamberts. Two neutral filters were also used giving illuminations of 3·9 and 0·39 millilamberts. Monochromatic filters (Kodak nos. 2, 4, 6, and 8 and a matched neutral filter) were also used, all at the same intensity of illumination. They were matched by measuring the intensity of illumination, using the various filters with a photo-electric cell having approximately the same range of sensitivity as the human eye and adjusting the position of the screen accordingly. Moving the screen farther or nearer to the lamp altered the size of the stimulus field. This was compensated for by altering the position of the subject's head so that the stimulus field maintained the same retinal projection. (ii) A Standard E.M.I. Electric Stroboscope. This was placed 10 cm. from the opal glass screen with a sheet of tracing paper over the glass further to diffuse the beam. This lamp gave an area of fairly uniform illumination, some 20 cm. across, so the subject was placed with the tip of his nose touching the lower limit of this area thus obtaining fairly uniform stimulation of all his visual field except the extreme periphery which was at a lower level of illumination. The intensity of illumination (at 20 c./s.) was 1·0 log foot lamberts at the centre of the screen.

Thirty-five subjects were used, mainly advanced psychology students or staff of the department, and, these, after some trial runs, were given the task of describing the visual phenomena evoked by the stimulus under the following set conditions. After 30 sec. of dark adaption both eyes were opened and exposed to the flicker for 1 min. and the phenomena described. Then both eyes were shut and covered for 30 sec. and any after-images described. Each eye was then exposed for 1 min., closed again for 30 sec. and then the other eye was exposed for 1 min. and closed again for 30 sec., the phenomena and after images being described throughout. Both eyes were then opened for 30 sec. and then one eye was closed for 1 min., opened again for 30 sec. and then the other eye was closed for 1 min. and then opened again and the phenomena described throughout. Thus the effects of stimulating both eyes together could be compared with the effect of stimulating each eye singly, and the development of the patterns in time, the effects of 30 sec. dark adaption and of suddenly closing one eye could be followed. In twelve cases eighteen such runs were done with the Aldis lamp (two runs at each of the three levels of illumination and at each of three rates of flicker 6, 12 and 18 c./s.) and six runs with the stroboscope (two at each of the three rates of flicker) and then two runs with each of the colour filters and the matched neutral filter at 12 c./s. In seventeen cases the first series was omitted and in five only the second series was done. In one case (the experimenter) a somewhat different procedure was followed. The first run in each pair was in the order given above and the second in the reverse order to avoid artefacts due to satiation, etc. The subjects were also encouraged to draw their patterns.

III. RESULTS

After the trial runs coherent descriptions of the phenomena were obtained and the effects were very constant for each individual over many months. The bright phase patterns will be described elsewhere and this paper will be concerned with a description of the quite different dark phase patterns and with an inquiry into their origins. This phenomenon was found to be considerably more complex than described by Brown and Gebhard and seven more or less clear-cut types may be recognized. Moreover, the main features distinguishing the two phases were the facts that the bright phase patterns were predominantly geometrical and flickered whereas the dark phase patterns predominantly showed no trace of geometrical form and never flickered. The seven types were as follows:

(i) *Amorphous*

This consisted of two colours, usually a red and a green, sometimes in more than one shade and most often rather unsaturated. These two swirled around in a very characteristic motion described as being just like oil on the surface of water, or sometimes the motion was of a thicker liquid like boiling mud or porridge. The following were recorded descriptions of this (each number corresponds to a different subject): (13) A very definite movement, oily, steady and very slow with several swirls, currents, whirlpools, spirals—like liquid stirred up...a slow oily motion in bright green and a fairly deep brown, very irregular splodges, and a large number of vortices, both clockwise and anti-clockwise.... It is a very odd effect indeed, very clear and marvellously pretty. (15) Hot molten earth heaving through the flicker—a volcano—lava in reddish brown—I can see it quite clearly and can describe it more satisfactorily than the other. (18) Random oily swirling motion. (35) A swift oily swirl like light petroleum on the surface of water. (Five subjects reported this type of pattern. In one case (9) there was a swirly oily streaky effect without actual movement.)

(ii) *Particulate*

Here the same oily swirling motion was obtained but small objects swirled rather than mere sheets of colour. These 'objects' were described as looking like (*a*) bacteria seen under a microscope or pond life or powder on a liquid surface (7 subjects); (*b*) small lines and squares described as 'streets and houses' swirling around (2 subjects); (*c*) an antheap (1 subject); (*d*) small pieces of snow-flake or pine cone (1 subject); or (*e*) a transition to class 7 *a*—a shoal of fish (1 subject).

Examples. (*a*) (21) Small things wriggling or swirling around, not flickering at all—lots of green bacteria things on a pinky-grey ground...(sometimes the particles were larger and were likened to curds and whey). (14) Powder on a liquid surface...grains of sand swirling around...a swirling granulated pattern. (6) Brown and green flecks swirling like a cloud...green splodges swirling about. (24) Pond life...germs under a microscope wriggling madly. (26) Lines and dots in a very peculiar smooth movement like a very densely populated slide of bacteria...a continuous rotary and horizontal movement with eddies. (28) A reddy-grey field with green bits writhing on it. (33) Red swirly patches of plankton in lines.

(*b*) (18) An aerial photo of a city with streets and blocks of houses. This starts moving around—some green and some red. It is not the same as the antheap—the colours are different and more positive; the things are bigger and better defined. (7) It is like looking at a city and seeing streets and blocks of houses, but all swirling about...it is like looking at London from a tremendous height and seeing the whole lot swirling about.... The elements are very small and very numerous.

(*c*) (18) The movement is like something vaguely familiar—like an antheap with random movement—a continuous sometimes swirling movement of little things.

(*d*) (13) I can see little bits like snowflakes under a microscope, almost three-dimensional. It is a very restful movement and a very deep red-brown almost maroon; the ground is a very deep green...not snowflakes, just needles; it is like watching the organization of crystals. (6) There are very pale orange flecks moving in the grey, swirling around; they look like frost patterns on a window, relatively clear.

(*e*) (16) Lovely tropical fish in a blue tank; a tank full of fish and bubbles. The tank is purple and the fish greenish. The fish are weaving in and out of the weeds. The tank is crammed with fish. They are fish-like shapes—angel-fish— but they have no details of fins or eyes.

(iii) *Stationary patterns*

These again were of various kinds: (*a*) Paisley patterns (2 subjects); (*b*) marble (2 subjects); ink-blots (2 subjects); (*d*) map, and (*e*) leaf patterns (both in 1 subject). In some cases the patterns were covered with a remarkable effect looking exactly like very clear rippling water.

Examples. (*a*) (15) The patterns are Paisley patterns in dark red and various browns and I am looking at them through clear deep water. (24) A Paisley pattern effect in red and green.

(*b*) (25) A flat marble wall all gently shimmering like the bottom of a swimming bath with water flowing over it—it is very solid and real. The marble has a variegated effect with mineral lines. (25) An effect like rough stucco or mineral in shades of brown with clear, clear water rippling gently over it. It is most impressively clear and distinct.

(*c*) (18) A rather messy inkblot kind of pattern. (33) A very complex design in several layers; a fixed ground just like an inkblot, then on this there is pond-life and the front layer is of separate rather large spherical objects.

(*d*) and (*e*) (11) (The 'leaf' pattern was made up of small to very small pieces of colour with *curved* lines between looking like leaves on black water)—like thousands of leaves with different coloured veins all in a heap. (This could swirl about as if on black water. The 'map' pattern was composed of sharp *straight* lines of various thickness (roads) with small sharply outlined blocks along them (houses) and largish blank areas (fields). The proportion of houses and fields varied. This might also swirl about or be quite stationary with the shallow water effect swimming over it.)

Note. Clearly 3 (*e*) and 2 (*b*) are the same pattern either stationary or swirling respectively. Similarly, as was noted above, pattern 2 (*e*) is transitory to pattern 7 (*a*). Patterns 3 (*a*), (*b*), (*c*) never swirled: patterns 2 (*a*), (*c*), (*d*), (*e*) only rarely did not.

(iv) *Watery patterns*

Examples have been given above of the clear water that could swirl or ripple over the surface of the class (3) patterns. This effect was extraordinarily vivid. A variant of this was an effect like the reflexion of rippled water in sunlight on to the undersurface of a bridge. These effects were found in 5 subjects.

(v) *Design patterns*

A curious variety was an ornamental design described as Victorian wall paper or a modern abstract design (3 subjects).

Examples. (25) A Victorian wall-paper with dark green bits of fern—A Victorian fern wall-paper—very deep and strangely distinct—these ferns were etched in deep green on a white surface. (15) There is a dark old-fashioned Victorian wall-paper under water—a repetitive pattern of dark heavy non-representational shapes arranged with some regularity. (18) A dark green ground with a terrific modern design for a wall-paper of black lines with knobs on the end forming triangles.

(vi) *Animate patterns*

An 'animate' effect was obtained only in one subject (25) and was of some living thing coupled with a strong emotive sense of 'eerieness'. The following is his vivid description of the experience: It was as if I was looking into very deep water and seeing deep green coral or a coiled octopus in dark green and white—very deep. This stuff streaming out from the centre (the bright phase patterns) is just surface thought, but this was a deep thought. It impressed me a good deal psychologically like Jung's image—it has intrinsic significance or archetypal quality. The marble is deep; the ferns are deeper, but this was miles underneath it all.... A very unusual effect. It started as centrifugal motion (a bright phase pattern), then the nucleus began to have an odd look and a little irregular snowflake began to form at 'marble' depth and this grew and filled the whole field with a salmon-pink ground bearing repeated *identical irregular* snowflakes. These then all melted into the impression of a THING! These small snowflakes melted into one large snowflake which became alive; it turned into a living creature—slightly eerie—like Quattermass (a prominent science fiction character).

(vii) *Scenes*

Fully formed scenes were of two kinds (*a*) a single scene of formed objects was predominant, and (*b*) there was a succession of different scenes.

(*a*) This was described by two subjects. The first (16) was the borderline case described also under 2 (*e*) of the fish patterns. These varied in complexity and when ill-formed fell clearly into the 2 (*e*) category of a particulate swirl. However, when they were well formed they developed good object quality and had fins or stripes and presented the appearance of a shoal of fish in an aquarium. The second subject (2) had a constantly recurring scene of vegetation, usually simple clumps of grass or leaves. The grasses were usually in two clumps growing up from the lower lateral halves of the field and arching over in front of the subject's apparent point of observation. They possessed good object quality with little heads of seeds. The leaves were usually arranged around the periphery of the field leaving a clear centre. The subject described this as if he were looking out of a tree through a gap in the foliage. Both the grass and the leaves swayed about naturally as if blown by the wind. More complex examples of these vegetation patterns were: steamy forests going up a mountain side with a very dirty brown river below... excellent grey vegetation—grey trees—aspens—and different kinds of trees and grass—very intricate and seen very clearly... a very deep indigo ground with very clear grasses on this in outline and trees on a bank, extremely clear cut and distinct, waving about in the moonlight.

(*b*) Two subjects (5, 18) reported a continual stream of images of fully formed scenes, usually of commonplace objects and events such as trains, cars, street scenes, harbours, animals, people, etc. One subject described them as being like a number of scenes in a badly cut film. Four other subjects (2, 12, 16, 17) experienced such scenes occasionally. In one case (2) they were usually of animals, for example an elephant, a pheasant, a beetle, etc. The scenes entirely lacked the aesthetic appeal of the mescaline phenomena and hypnagogic imagery.

IV. GENERAL FEATURES OF THE DARK PHASE PATTERNS

(i) They were much more clear cut, more definite and easier to describe than the bright phase patterns.

(ii) The subjects in general found them very interesting, even fascinating, and one even expressed the conviction that they possessed a power of addiction over the subject as one developed an active desire to go on looking at them.

(iii) The patterns usually occurred on the nasal side of the field (the side of the externally dark eye) and then swept across the whole field from time to time except for the extreme temporal arc which they did not cover. They usually alternated with the bright phase patterns in a manner which reminded many subjects of typical retinal rivalry. They could also co-exist with the bright phase patterns in which case the former appeared as a solid wall behind a flickering diaphanous curtain of the bright phase.

(iv) They were only seen with single eye illumination and never with both eyes subjected to the flickering field.

(v) The dark phase patterns frequently continued as clear-cut positive after-images which the bright phase paterns never did.

(vi) Many subjects expressed the conviction (not suggested by the experimenter) that the dark phase patterns were being seen with, or were 'in', the externally dark eye. Their reasons were that the patterns alternated with the bright phase patterns in a manner that reminded them strongly of retinal rivalry; that they did not cover the temporal arc of the field and were mostly on the nasal side.

Table 1

Proportions of dark phase colours produced by different coloured lights. Significant differences from the white control are printed in bold type. *D*, dark phase colour; *A*, Aldis lamp; *S*, strobe; *W*, Aldis control matched with monochromatic lights for intensity; *R*, red; *Y*, yellow; *G*, green; *B*, blue; *M*, browns, buffs, etc.; *N*, black and grey; *O*, number of reports counted. The proportions in the case of the dark phase rows *RYGB* is of the total *RYGB*. In the case of *MNO* it is of the total *RYGBMNO*. Column *A* is based on data from 5 subjects, S from 19 and *W*→*B* from 18

D	*A*	*S*	*W*	*R*	*Y*	*G*	*B*
R	0·19	0·42	0·28	**0·18**	0·28	**0·46**	0·21
Y	0·07	0·05	0·04	0·07	0·05	0·09	**0·21**
G	0·59	0·48	0·46	0·43	0·40	**0·19**	0·50
B	0·15	0·05	0·22	0·32	0·27	0·26	**0·08**
M	0·16	0·15	0·03	0·06	0·03	0·15	0·13
N	0·06	0·13	0·13	0·24	0·21	0·20	0·11
O	114	179	92	122·5	116	130·5	139

(vii) The colours of the dark phase patterns reported were shades of red, orange, yellow, green, blue, violet, purple, browns and clay and mud colours, and black and grey. These were always in pairs: sometimes one member of the pair was in two shades or two states of dark and pale. Some quantitative measure of the relative frequency of these was obtained as follows. All reports of colours of the dark phase patterns in all the runs using white light in all the subjects were counted, tabulated, and scored by counting each report of a primary colour red, yellow, green or blue as 1; each report of a mixed colour such as orange, yellow-green, green-blue or purple counted ½ for each of the constituent colours; and grey or browned colours counted ½ to grey (or brown) and ½ to the particular colour concerned. The results are expressed in Table 1. These show that green and red predominate over yellow and blue. The significance of these proportions was calculated and showed $w = 12$ standard errors which is a very significant result. The relative proportions of red and green were also assessed and came to $w = 3{\cdot}8$ standard errors which is significant. All the values for white light were pooled for this calculation.

(viii)

Subjects showing one type of pattern	7
Subjects showing two types of pattern	6
Subjects showing three types of pattern	5
Subjects showing four types of pattern	0
Subjects showing five types of pattern	0
Subjects showing six types of pattern	1
Total subjects showing dark phase patterns	19 out of 34

(ix) Some of the dark phase patterns may also be obtained as follows: cover one eye and observe a fairly brightly illuminated white surface (from a steady source), after some minutes I have been able to observe typical 'rippling water' patterns in this manner. A similar effect is noted by Nagel as a footnote in Helmholtz (1924).

V. FURTHER EXPERIMENTAL INVESTIGATION

(i) Experiments were conducted to see if any further evidence could be obtained as to the location of the dark phase patterns and to test Brown and Gebhard's hypothesis that they arise in the neurones belonging to the externally dark eye. If these patterns do arise in the optic system of the externally dark eye (the term 'optic system' refers to all parts of the central nervous system connected particularly to one eye), then stimulation of this eye by some means might specifically affect these patterns; whereas, if they arise in the system of the illuminated eye, stimulation of the externally dark eye should have no particular effect. Brown and Gebhard used this method by flashing a weak steady light into the externally dark eye and noted that this always brought on the dark phase patterns. However, it is difficult to be certain without quantitative timing experiments whether such a change might not represent the normal fluctuation between the two phases. In the experiments reported here a quantitative method was used.

(*a*) *Qualitative investigation.* At first the effect of pressure on the externally dark eye was tried. This did not prove satisfactory as the degree of pressure could not be controlled properly. Several reports were obtained, however, that the pressure clearly immobilized the swirling patterns.

(*b*) *Quantitative investigations.* (i) Materials and methods. A mask was made to fit over the upper part of the face covering both eyes. In front of the right eye was a circular piece of frosted glass 5 cm. in diameter, 1 cm. from the eye. The unfiltered beam from the Aldis was projected on to this through the rotating episcotister. The other eye was covered with a similar piece of glass behind which was a 6-volt bulb connected to a variable resistance so that a dim uniform steady light could be applied to the left eye. Four levels of illumination were used *O* (none); *L*, 0·02; *LM*, 0·2; *HM*, 1·0; and *H*, 3·0 millilamberts. Wearing the mask the subject saw a flickering field with the right eye and a steady field, or only darkness with the left eye. These two fields overlapped in the subjective visual field and the subject was instructed to look straight ahead and to report when the dark phase patterns came and went. Each subject was quite familiar with his own dark phase patterns and could usually report these end-points without hesitation. Some difficulty was experienced with the brightest light. In this case the defining criterion of a non-flickering pattern was used. Five subjects were used with six runs for each subject. Each run consisted of eighteen consecutive 1 min. exposures during which the subject merely said 'there' and 'gone' to mark the limits of the definite appearance of the dark phase patterns in his field. There was 30 sec. rest between each exposure during which the subject kept his hand over the right eye-glass during most of the experiments. The order of exposure to the different intensities of illumination was varied as shown in Table 2 to avoid fatigue effects producing artefacts.

(ii) *Results.* The results are shown in Table 3. These results seem to indicate clearly that a weak light (*L* was hardly visible) in the externally dark eye increases the proportion of the dark phase whereas a brighter light decreases it. However, as bright lights (H) in the left eye inhibited the development of the dark phase patterns it appeared possible that this effect might last longer than the 30 sec. rest period of darkness between each exposure. Examination of the order of exposures showed that twenty-two L exposures

Table 2

Showing the ordering of the stimuli to the externally dark eye (see text). Runs 2, 4, and 6 reverse the order of A and B. O, off; L, low; LM, low medium; HM, high medium; H, high

A 1	*A* 3	*A* 5	*B* 1	*B* 3	*B* 5
O	*LM*	*HM*	*O*	*L*	*H*
LM	*HM*	*O*	*L*	*H*	*O*
HM	*O*	*LM*	*H*	*O*	*L*
LM	*HM*	*O*	*L*	*H*	*O*
HM	*O*	*LM*	*H*	*O*	*L*
O	*LM*	*HM*	*O*	*L*	*H*
HM	*O*	*LM*	*H*	*O*	*L*
O	*LM*	*HM*	*O*	*L*	*H*
LM	*HM*	*O*	*L*	*H*	*O*

Table 3

Showing the mean number of seconds in a minute occupied by the dark phase patterns at different strengths of illumination of the externally dark eye in 5 subjects

	Illumination in externally dark eye				
Subject	*O*	*L*	*LM*	*HM*	*H*
Z	19·9	32·7	31·1	20·3	6·0
S	16·8	37·4	36·2	25·2	12·0
P	41·6	41·8	28·4	0·0	1·2
B	21·7	43·1	46·7	18·8	3·6
M	35·0	39·6	37·0	22·8	14·7
Mean	27·0	38·9	35·9	17·4	7·5

Table 4

Showing: Column 2. The mean time of the initial report of the dark phase patterns, column 3 the mean duration of the dark phase patterns per minute, column 4 gives the difference between two collated rows and column 5 gives the cumulative difference where relevant. Column 6 gives the number of reports of runs with dark phase component of less than 30 sec. and column 7 the same of 30–60 sec. Column 1 gives the state of illumination of the externally dark eye as to the present run and the previous one (see text)

1	2	3	4	5	6	7
O after *L*	13·3	32·8	—	—	11	11
O after *LM*	10·0	30·1	—2·7	—	13	12
O after *HM*	18·4	26·4	—3·7	—6·4	38	20
O after *H*	19·4	23·7	—2·7	—9·1	39	19
L after *O*	7·1	41·3	—	—	7	51
L after *H*	10·9	32·1	—9·2	—	7	15
LM after *O*	9·7	35·9	—	—	20	40
LM after *HM*	9·4	37·6	+1·7	—	6	14
HM after *O*	28·6	18·0	—	—	20	5
HM after *LM*	35·2	21·1	+2·1	—	44	15
H after *O*	39·6	9·7	—	—	25	1
H after *L*	43·8	6·7	—3·0	—	55	4

followed H exposures and fifty-eight L exposures followed O exposures whereas fifty-eight O exposures followed H exposures and twenty-two O exposures followed L exposures. Such a delayed effect, if present, would therefore lead to a spuriously low figure for the O exposures, for the O exposures would be more subject (than would the L exposures) to the delayed inhibition of the preceding H exposures. The runs were therefore analysed with respect to the level of the preceding exposure and these results are shown in Table 4.

This shows that exposure to a bright light in the left eye will depress the development of the dark phase patterns on stimulation of the right eye with the flickering stimulus 30 sec. later even with an intervening period of darkness lasting 30 sec. for both eyes. Notice in Table 4 the progressive effect of the inhibition of preliminary *L*, *LM*, *HM*, and *H* exposures and also that the effect is only obtained with subsequent *O* and *L* exposures, presumably because subsequent brighter lights swamp the effect in some way. This effect is also shown by the time of the start of the dark phase per run (column 2 in Table 4). The significance of these ratios is assessed in columns 6 and 7 of Table 4 which shows the distribution of the number of exposures showing more than 30 sec. dark phase activity per minute and the number showing less as between the various conditions. The significance of the relative number of below thirty and above thirty runs obtained in *O*, *L* and *LM* exposures following *O*, *L* and *LM* exposures compared with those following *HM* and *H* exposures is given by $\chi^2 = 22$ for one degree of freedom, which is very significant. Notice further from Table 4 that, whereas some of the difference between the *O* and the *L* figures in Table 3 must be due to this artefact (because more *O*'s and fewer *L*'s are preceded by the inhibiting *H* exposures), not all of it need be, for the figure for *O* after *L* only just overlaps the figure for *L* after *O* and the figure for *L* after *H* is considerably above that for *O* after *H*. Further experiments are needed to clarify this point. The figure for *LM* in Table 3 is not affected by the analysis in Table 4 as there is not a significant difference between the figures for *LM* after *O* and *LM* after *HM*. However, the delayed inhibition of the dark phase patterns by a previous exposure of the externally dark eye to a bright light is itself of great interest.

(2) Experiments were also conducted to see if the colour of the stimulating light affected the colour of the dark phase patterns. The colours were scored as described for white light above and the results are also shown in Table 1. This shows that monochromatic red, green and blue depress the reproduction of the same colour in the dark phase patterns and that blue and green enhance the production of their complementary colours (as compared with white light). The colours produced by yellow were very similar to those produced by the white control. The significance of the reduction of the same colour in the dark phase patterns by monochromatic light (red, green and blue) was assessed by χ^2 by comparing the number of reports of dark phase colours for each stimulus colour as compared with the number of reports of each dark phase colour with the white control stimulus. The value for χ^2 was 15 with 2 degrees of freedom which is significant at the 0·1 % level. χ^2 was also estimated in the same manner for the increase in the amount of complementary dark phase colour produced by stimulus with green and blue monochromatic lights. The value was 146 with one degree of freedom, which is very significant. It will also be noted in Table 1 that the yellow light of the Aldis projector produces more blue than the blue light of the stroboscope, thus keeping to the rule that a monochromatic stimulus light will tend to suppress the production of the same colour in the dark phase.

VI. DISCUSSION

The evidence that the dark phase patterns arise in the optic system of the externally dark eye is as follows:

(i) They only occur with uniocular stimulation and never with binocular stimulation.

(ii) They are located mainly on the nasal side of the field and do not cover the extreme temporal arc.

(iii) They alternate with the bright phase patterns (which certainly arise in the stimulated eye) in a manner which has reminded many experienced observers in both Brown and Gebhard's series and our own of retinal rivalry.

(iv) They are probably increased by shining a dim light (0·02–0·2 millilamberts) into the unstimulated eye and are inhibited by shining a brighter light into this eye (1·0–3·0 millilamberts). Moreover, this latter inhibition is still effective in the period from 30 to 90 sec. after exposure.

These facts may most easily be explained by the hypothesis presented above. On the other hand, if we suppose that the dark phase patterns are produced in the optic system of the stimulated eye, one would have expected that they would have been as easily visible with both eyes stimulated as with only one. Furthermore the nasal location of the patterns would remain unexplained and the results reported in (iv) above could only be explained by supposing that the traces of previous activity in the optic system of one eye could thus selectively affect the subsequent development of the dark phase patterns in the optic system of the *other* eye. This would appear to be a less plausible hypothesis than the alternative—that traces of previous activity in one optic system affect the subsequent development of the dark phase patterns in the *same* system. Thus there is good, though not yet conclusive, evidence that the dark phase patterns originate in the optic system of the externally dark eye.

The next problem to decide is whether the interaction occurs at (i) a transorbital level, (ii) a cortical level, or (iii) by means of the corticofugal fibres to the retina. The first hypothesis is unsatisfactory in that it does not explain why the dark phase patterns should be so clear cut, definite and so different from the bright phase patterns, for one would expect them merely to be faint replicas of the bright phase patterns; nor does it explain why they should have lost all trace of flicker. Nor does it explain why the dark phase patterns characteristically develop after some seconds' stimulation, unless one supposes that the unexposed retina must dark adapt for some time before it can react to the dim illumination filtering across from the other orbit. But this is incompatible with the fact that dim illumination of the externally dark eye increases the amount of the dark phase at the same time as decreasing the dark adaption of the eye. The explanation most conversant with all the facts is that this is a cortical effect; the neurones connected to the right retina must surely be closely mingled with the neurones connected to the equivalent location on the left retina. It is possible that there may be connexions between the neurones connected to the right eye (*R* eye neurones) and those connected to the left eye (*L* eye neurones) in the cortex. These connexions might subserve the construction of the unified binocular visual field and would only normally be activated when both sets of neurones were activated from their sense-organs. But in the case of this highly unusual intermittent stimulus activity of the *R* eye neurones might leak over, as it were, into the inactive *L* eye neurones along these paths (when the *R* eye is stimulated) and produce these perceptions of the swirling, non-flickering dark phase patterns. Another explanation may be based on Gregory's hypothesis (personal communication, 1957) that there is a cortical filter between the *R* eye neurones and the *L* eye neurones which cuts out all signals but lets through noise. In which case, when the *R* eye neurones, for example, are activated by the flickering stimulus any noise generated in this process would be conducted to the *L* eye neurones, and there, as these neurones are not being activated from their normal source, they would give rise to aberrant perceptions. It is of interest here to note that the appearance of the predominant dark phase patterns is very similar to that of visual noise presented on a television screen. The light dots of the latter swarm about, under certain conditions, in a manner very reminiscent of the 'bacteria' or 'pond-life' of the dark phase patterns, and, moreover, under certain conditions, effects very like the rippling water patterns of the dark phase may be observed. The dark phase patterns may, in fact, be an internal visualisation of one's own visual noise.

The third hypothesis, that corticofugal fibres going to the externally dark eye may also be responsible for the phenomenon, cannot be excluded, but the second hypothesis of cortical spread is certainly simpler. However, Muller-Limmroth's finding (1955) that light shone in one eye can produce slow potential changes in the other eye—the nervous pathway going via the lateral geniculate body—may be relevant here. Hypotheses (i) and (iii) could be eliminated if the phenomenon could be obtained in a person with only one eye. I have examined two such cases, but they did not possess sufficient skill to report their patterns unambiguously. Further experiments in this line are being conducted.

I am most grateful to Professor O. L. Zangwill for his continued support and interest in these experiments, to Mr E. G. Chambers for advice on the statistics and I should like to record my debt to Professor Heinrich Klüver for developing the idea that the form constants of hallucinations represents a worthwhile field of study.

REFERENCES

BROWN, C. R. & GEBHARD, J. W. (1948). Visual field articulation in the absence of spatial stimulus gradients. *J. exp. Psychol.* **38**, 188–200.

GREY, WALTER W. (1950). In *Perspectives in Neuropsychiatry* (ed. D. Richter). London.

HELMHOLTZ, H. VON (1924). *Physiological Optics*, vol. II (ed. James Southall). The Optical Society of America, p. 8.

MULLER-LIMMROTH, H. W. (1955). Der Einfluss der Belichtung auf das Bestandpotential des Froschauges. *Z. Biol.* **107**, 216–40.

PURKINJE, J. E. (1823). *Beobachtungen und Versuche für Physiologie der Sinne. Beiträge zur Kenntnis des Sehens in subjectiver Hinsicht.* Bd. I, Prag.

(*Manuscript received* 23 *September* 1957)

THE TOLMAN AND HONZIK INSIGHT SITUATION

By W. S. ANTHONY

Department of Psychology, Durham University

It is argued that an explanation in terms of processes simpler than insight would account for almost all the evidence relating to this situation, but that a result obtained by Evans is a notable exception, and is particularly favourable to the supposition that insight is involved and also to an 'expectancy' theory such as that of Deutsch.

I. Introduction

Deutsch (1956) has put forward a theory of insight, reasoning, and latent learning, which explains, *inter alia*, the 'result' of experiments on the Tolman & Honzik (1930) situation designed to test insight in rats. It is characteristic of Deutsch's (1953, 1956) theory as a whole, that motivated behaviour is regarded as a series of approaches to stimulus-situations. In so far as it relates to the Tolman & Honzik situation (which will be described below) the theory may be summarized as follows. A particular feature in the environment (such features are here called 'cues') activates a particular 'analyser' (a receptor-mechanism). In learning, analysers become arranged in the order in which they are activated, each becoming attached to one 'link' in a row of links. The first link in a row is sensitive to a particular internal state such as hunger, so that when the organism is hungry this link becomes excited, and transmits excitation to successive links in the row. The analyser attached to the second link will be sensitive to food, and analysers attached to succeeding links will be sensitive to cues which have been perceived at progressively longer periods of time before the perception of food. When its analyser is activated, an excited link stops transmitting excitation to succeeding links, and instead takes over control of behaviour, making the organism approach the activating cue. In this way an animal in a familiar maze progresses to the goal; for when a cue nearer the goal is perceived, the analyser it activates will be attached to a link nearer the source of excitation; this link will deprive all succeeding links of excitation, and the animal will approach the newly perceived cue. If a path is blocked, so that the approach to one cue does not lead to the perception of the next, the threshold of excitation of the controlling link will gradually rise, depriving it of control. Other paths in the maze will be represented by other sets of analysers attached to parallel rows of links; where paths meet, the analysers sensitive to the junction will be attached to one link in each row, and will transmit excitation between rows. As excitation is transmitted along a row its intensity diminishes. If, as may be at a choice point, links are 'competing' for control of behaviour, the link receiving most excitation will exercise the control.

From this theory it would be predicted that of two or more paths to the goal, the shortest will normally be taken, but that if the animal finds a block in the shortest path it will take a path (if there is one) which avoids the block, and will not take a path which leads back to the block. Thus Deutsch predicts success in the Tolman & Honzik insight situation, as may be seen from the description of the situation which is given below.

It is the purpose of the present paper to review briefly the relevant evidence. It will be argued that a great deal of the (apparently conflicting) evidence strongly suggests an explanation which is in terms of processes of a lower order than insight and not in accordance with Deutsch's treatment; but it will also be argued that there is one result, viz. Evans (1936), which is especially favourable to Deutsch's theory or any other 'expectancy' type theory. Deutsch (1956), Munn (1950), Osgood (1953), and Thorpe (1956) do not refer to Evans's result in their reviews of the situation; Caldwell & Jones (1954), Helson (1951), and Hilgard (1956) refer to it, only stressing, however, its negative aspect.

II. DESCRIPTION OF THE SITUATION

The general plan of the situation may be seen in Fig. 1. The subject is kept on a hunger schedule, and is trained to run for food on the maze shown; he is encouraged to acquire a preference for path 1 over paths 2 and 3, and for path 2 over path 3. In the test situation he runs up path 1 and finds a block beyond the point where paths 1 and 2 converge (hereinafter called the 'convergence point' *con.*). The block renders both paths 1 and 2 useless, but path 3 is still available. The rat has to retrace path 1 (there is a non-retracing gate at the end of path 2), until it reaches the choice point (*cho.*), where it chooses between paths 2 and 3. It is significant, however, that in Evans's experiment the rat does not retrace path 1, but is picked up and later replaced at the starting-point to make another choice from there. In any case path 2 is the stupid choice in this situation, and path 3 the correct one, and *prima facie* it seems that if the rat chooses path 3 it is behaving insightfully, since in the ordinary way it would prefer path 2. It is, of course, the first

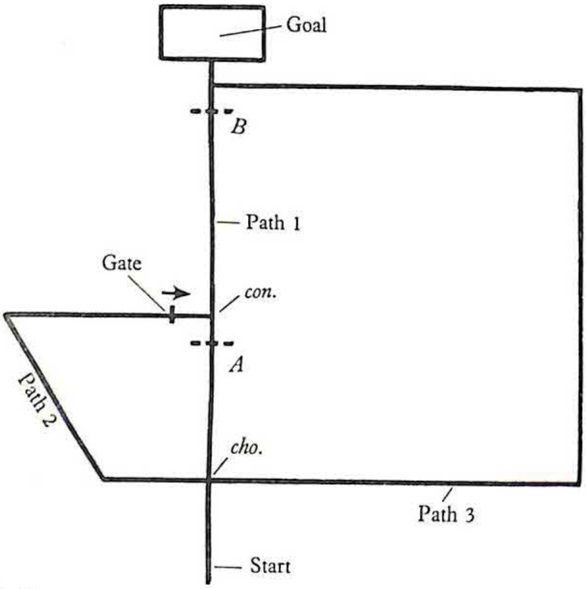


Fig. 1. General plan of the Tolman & Honzik situation. The shortest path to the goal is path 1; the next shortest is path 2; the longest is path 3. *con.*, 'convergence point' (of paths 1 and 2); *cho.*, choice point. The gate near the end of path 2 only permits passage in the direction of the arrow, i.e. towards the convergence point; it will not allow the animal to enter path 2 from the convergence point. If path 1 is blocked at the entrance, or between the entrance and the convergence point, e.g. at '*A*', then path 2 is the best path. In the test situation, after the subject has been given experience of the apparatus, he finds path 1 blocked beyond the convergence point, e.g. at '*B*', for the first time. Except in Evans's experiment, the subject then retraces to the choice point and has to choose between paths 2 and 3. Path 3 is now the only correct choice.

test trial which is important, since later good performance could be attributed to simple learning.

Deutsch (1956) only deals with the case where the animal is removed and replaced at the start. But his theory would also predict success in the retracing procedure which was used in the main body of the experiments. *In the test* situation, the rat is first drawn goalwards down path 1 by the high-intensity excitation passing down the row of links representing path 1. But when the rat comes to the block, the controlling link loses its excitability; so control of behaviour is taken over by a lower-intensity excitation which is transmitted via the row representing path 3, across the common analyser sensitive to the choice point, and back up the row representing path 1, until it reaches the link previous to the link whose threshold of excitation has risen. Since the low-intensity excitation is being transmitted in the opposite direction to that of the high-intensity excitation which is now inhibited, the rat will now be drawn back up path 1 to the choice point, and into path 3.

III. THE EVIDENCE AND AN EXPLANATION

A number of experiments have been done on the Tolman & Honzik situation; the evidence, with respect to the rat's choice between paths 2 and 3 on the first test trial, is given in Table 1. A 'positive' result means that on the first test trial the subjects mostly chose path 3; a 'negative' result means that they did *not* mostly choose path 3 on the first test trial.

Table 1. *The rats' choice on the first test trial*

(If over half the subjects in an experiment chose path 3 on the first test trial, the result is here called 'positive'; other results are called 'negative'. The exact proportions are given in the third column. It is maintained in the text of this paper that Evans's experiment is of particular importance.)

	Result	
Experiment	Positive or negative	Proportion choosing path 3
Tolman & Honzik (1930)		
Exp. II	−ve	5 out of 11
Exp. III, group B	+ve	7 out of 10
Keller & Hill (1936)		
Part I	+ve	6 out of 7
Part II	−ve	1 out of 7
Evans (1936)	+ve	23 out of 29
Kuo (1937)		
Series 1		
Exp. II (unblindfolded rats), group 2	+ve	9 out of 10
Exp. IV (blindfolded rats), group 2	+ve	7 out of 10
Series 2		
Exp. II (unblindfolded rats), group 2	−ve	3 out of 10
Exp. IV (blindfolded rats), group 2	+ve	8 out of 10
Dove & Thompson (1943)		
Exp. A	−ve	2 out of 11
Exp. B	−ve	0 out of 26
Exp. C	−ve	0 out of 14
Caldwell & Jones (1954)	+ve	19 out of 21

Not all of Kuo's results appear in the Table because his curious procedure detracts from their relevance for our purposes. His subjects were divided into two groups (1 and 2); the results given here are for group 2, which in training was given a forced bias towards

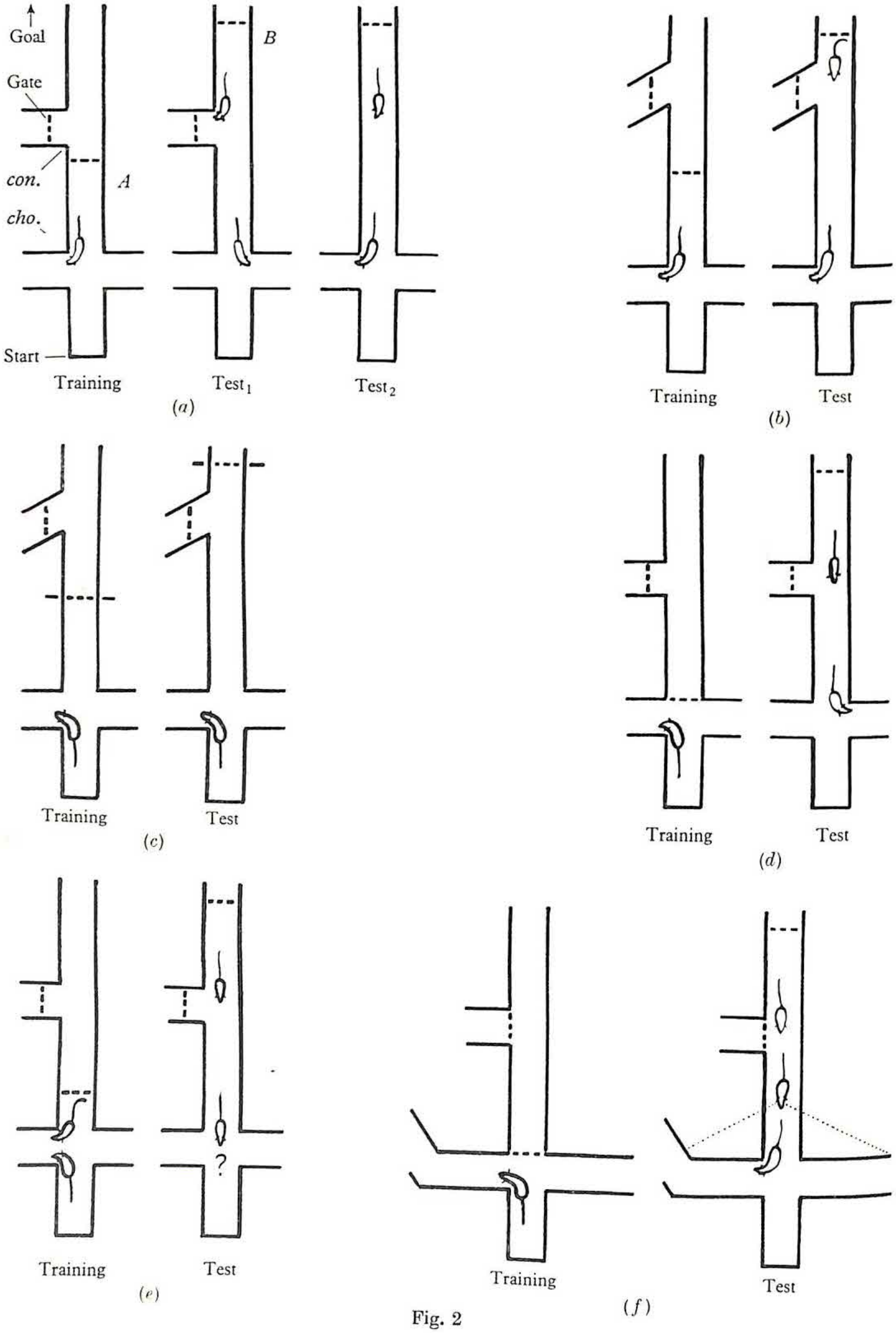

Goal
Gate
con.
cho.
Start
A
B
Training
Test$_1$
Test$_2$
(a)
Training
Test
(b)
Training
Test
(c)
Training
Test
(d)
?
Training
Test
(e)
Training
Test
(f)

Fig. 2

path 2; group 1 in training was given a forced bias towards path 3, and their results are opposite to those obtained for group 2.

The positive result for Tolman & Honzik's Exp. III, group A, subjects, is also omitted through its irrelevance. Tolman & Honzik's Exp. I (negative result) was a repeat, with modifications, of Hsiao's (1929) experiments. Hsiao's results on Part I of his work were negative, but on Part II they were positive. However, an examination of Hsiao's apparatus, procedure, and 'behaviour graphs' allows of an interpretation of his positive results in terms of simple extinction. The particular objection does not apply to Tolman & Honzik's and subsequent experiments.

The detailed observations of Keller & Hill (1936), in particular, point to an explanation in 'simple' terms which covers much of the evidence, with the notable exception of Evans's experiment. The way in which such an explanation works is indicated in Fig. 2.

Fig. 2 (*a–f*) shows a 'simple' interpretation of the rats' choice on the first test trial of the Tolman & Honzik situation, according to the various experimental conditions. The diagrams represent the part of the maze which, on the test trial, the rat experiences before making his critical choice, i.e. the start, the choice point (*cho.*), the convergence point (*con.*), and the rest of path 1 up to the test-position of the block. Two basic assumptions are made: (i) that in an enclosed maze a rat is liable to mistake its orientation, (ii) but not on an elevated maze, if its vision is normal. Each of the parts (*a–f*) of the figure represents one or more different experiments, as follows. (*a*) Tolman & Honzik III, group B, Keller & Hill, elevated maze. The rat has been trained to turn right after retracing from a block in path 1; in the test, he tries to turn right at the convergence point, is prevented by the one-way gate (or, in Keller & Hill, by hand) from proceeding, and turns left at the choice point (positive result due to extinction of right-turning tendency). Keller & Hill, in a second test, removed the end of path 2; here the rat is 'puzzled' at the former convergence point but, the full response not being aroused, turns right at the choice point (negative result from lack of extinction). (*b*) Dove & Thompson B, elevated maze. The convergence point is made more distinctive both in form and in spatial relations, so the right-turning tendency is still intact at the choice point (negative result). In their Exp. A, a repeat of Tolman & Honzik III, group B, we would expect a positive result; it was in fact negative on the first test, but was distinctly more positive than B or C on later tests. (*c*) Dove & Thompson C, elevated maze. The block is made very visible; in training, the rat turns left at the choice point without approaching the block; in the test he turns left as before (negative result, in later trials distinctly more negative than A or B). (*d*) Kuo Series 1, Exps. II and IV, group 2; Caldwell & Jones; enclosed maze. Rat is trained to turn left into path 2; on test trial rat is disoriented and turns left as before (positive result). This explanation also accounts for Kuo's other results, including that for the blindfolded rats on the elevated maze (Series 2, Exp. IV). (*e*) Tolman & Honzik II, enclosed maze. In training, rat sometimes approaches block (set back a little way) and sometimes turns immediately; hence no definite turning-response is established (negative result). (*f*) Kuo Series 2, Exp. II, group 2, elevated maze. Rat is trained to turn left into path 2; in test, becomes correctly oriented (undisturbed by convergence point, since it has not been trained in that direction), recognizes path 2 as the usual path, and enters (negative result). Residue of unexplained results: Dove & Thompson A (negative), Evans (positive); Evans's experiment is not pictured here, but is described in detail in the text.

Dove & Thompson (1943) carried out three experiments, 'A', 'B' and 'C'. Exp. 'A' was a repeat of Tolman & Honzik's Exp. III, group B, and was on maze I, a copy of the maze used in that experiment. Independently of Keller & Hill, Dove & Thompson state the 'simple' explanation as it applies to Tolman & Honzik's Exp. III, group B. Their Exp. 'B', on maze II, in which the convergence point was more distinct from the choice point than in maze I, is a test of the explanation; the results partially confirm it (not wholly, owing to their unexpected result on Exp. 'A'). In Exp. 'C', maze II was again used; in this experiment the block at *A* in training was made easily visible, so that the rat did not enter path 1 when it was blocked; for the test runs the block was moved back to *B*, and the rat again made his choice between paths 2 and 3 without entering path 1.

The authors write, 'all choices were made while the animal was facing the food box enabling him to "see" the entire situation, an important factor in our estimation'. The present writer does not agree with their suggestion that insight should be regarded as a kind of sight, and thinks it more reasonable to encourage the rat to approach the block rather than to try to assess the situation at a distance. It would be interesting, however, to know whether, in training, blocking the paths has any disturbing effect on the animal; Deutsch and Anthony (1958) report a disturbance produced by the use of blocks in training on a two-choice maze.

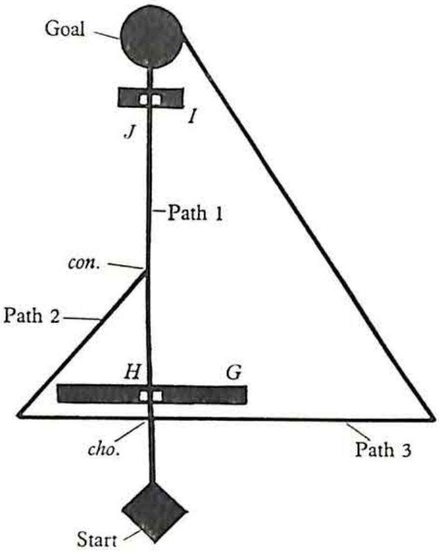


Fig. 3. Elevated maze with screens, used by Evans. *cho.*, choice point; *con.*, convergence point. The screens, *G* and *I*, are sheets of galvanized iron, each 10 in. high, and contain the doors *H* and *J* respectively. The doors are 6 in. square, and are hinged at the top. For the experimental and first control group, the position of the screens was as shown; but for the block control group, screen *I* was placed about half-way between the convergence point and screen *G*. For both the experimental and block control groups, door *H* is closed for the first stage of training only, and door *J* is closed for the third stage of training only. Thus, in the third stage of training (the test situation), path 3 is the only open path for the experimental group, and path 2 is the shorter open path for the block control group. For the first control group, door *H* was always closed and door *J* always open.

The explanation indicated in Fig. 2 does not apply to Evans's result. His experiment was on an elevated maze with screens, as shown in Fig. 3. There were three groups of subjects: an experimental group and two control groups. The experimental subjects were first trained with path 1 blocked near the choice point, so that they acquired a consistent preference for path 2 over path 3. Then they were given a second stage of training, with all three paths open, and acquired a consistent preference for path 1. Finally, they were given a third series of trials, in which path 1 was blocked between the convergence point and the goal. In this series, on the first trial, all subjects chose path 1 and approached the block; then, *on retracing path* 1, *they were removed* before they reached the convergence point, and after 10 or 15 minutes were replaced at the starting-place, and allowed to choose again between the three paths. Any subject which again chose path 1 was removed and given another trial, as before. This procedure was continued until every subject had chosen either path 2 or path 3. Out of twenty-nine subjects, twenty-three chose correctly after first abandoning path 1, and only six chose path 2.

Simultaneously with the procedure for the experimental group, a control group was trained on the maze, for which path 1 was always blocked near the choice point; for this group, no great change-over of preference occurred at any stage of the procedure. A second control group, the block control group, was trained in the same way as the experimental group, except that the second stage of training lasted for 15 days instead of for 25 days, and that on the third stage, path 1 was blocked between the convergence point and the choice point instead of between the convergence point and the goal. For this group, out of twenty-six subjects, fourteen chose path 2 (the correct path) after first abandoning path 1, and twelve chose path 3, even though, as for the experimental group, all the subjects had previously acquired a consistent preference for path 2 over path 3.

IV. CONCLUSION

In Evans's experiment, both the experimental and the block control groups had acquired consistent preferences for path 2. In the test situation, for the experimental group, paths 1 and 2 were closed, and twenty-three subjects (80%) chose path 3; for the block control group, only path 1 was closed, but only fourteen (54%) chose path 2. This result suggests that for both groups there was a tendency in the test situation to take the previously unpreferred path; this led the author to stress the 'flexibility of established habit'. However, our attention should not be distracted from his other conclusion, that the results constitute a positive result in the Tolman & Honzik situation. Deutsch would not predict that so many (46%) of the block control group would choose path 3, but he would predict the difference in performance of the experimental and block control groups. Furthermore, this difference cannot be explained by the 'turning-response' method. For since the animals were picked up on their retracing of path 1, and placed again at the start to make their critical choice, none of the 'turning-response' considerations apply; there is no 'simple' explanation of the difference in performance of the experimental and block control groups on the lines of our previous explanations.

This does not mean that our previous explanations have been invalidated, for the experiments of Kuo and of Keller & Hill practically force us to conclude that 'turning-response' considerations apply in some situations at least, and Deutsch has made no attempt (as we have done) to explain the negative results. The principal grounds for regarding any positive results on the Tolman & Honzik situation as support for Deutsch or as evidence for insight, seem now to rest on the Evans experiment. Deutsch & Clarkson (in an experiment not yet reported), in a rather similar situation, using an enclosed maze, have obtained positive results which also appear to escape a 'simple' explanation.

At the time of writing, the author held a research training scholarship under the Medical Research Council at the Institute of Experimental Psychology, Oxford University.

REFERENCES

CALDWELL, W. E. & JONES, H. B. (1954). Some positive results on a modified Tolman and Honzik insight maze. *J. Comp. Physiol. Psychol.* **47**, 416–18.

DEUTSCH, J. A. (1953). A new type of behaviour theory. *Brit. J. Psychol.* **44**, 304–17.

DEUTSCH, J. A. (1956). A theory of insight, reasoning and latent learning. *Brit. J. Psychol.* **47**, 115–25.

DEUTSCH, J. A. & ANTHONY, W. S. (1958). Blocking the incorrect alley in a two-drive learning situation. *Quart. J. Exp. Psychol.* **10**, 22–8.

Dove, C. C. & Thompson, M. E. (1943). Some studies on 'insight' in white rats. *J. Genet. Psychol.* **63**, 235–45.

Evans, S. (1936). Flexibility of established habit. *J. Gen. Psychol.* **14**, 177–200.

Helson, H. (ed.) (1951). *Theoretical Foundations of Psychology.* New York: Van Nostrand.

Hilgard, E. R. (1956). *Theories of Learning.* New York: Appleton-Century-Crofts.

Hsiao, H. H. (1929). An experimental study of the rat's 'insight' within a spatial complex. *Univ. Calif. Publ. Psychol.* **4**, 57–70.

Keller, F. S. & Hill, L. M., Jr. (1936). Another 'insight' experiment. *J. Genet. Psychol.* **48**, 485–89.

Kuo, Z. Y. (1937). Forced movement or insight? *Univ. Calif. Publ. Psychol.* **6**, 169–88.

Munn, N. L. (1950). *Handbook of Psychological Research on the Rat.* Boston: Houghton Mifflin.

Osgood, C. E. (1953). *Method and Theory in Experimental Psychology.* New York: Oxford University Press.

Thorpe, W. H. (1956). *Learning and Instinct in Animals.* London: Methuen.

Tolman, E. C. & Honzik, C. H. (1930). 'Insight' in rats. *Univ. Calif. Publ. Psychol.* **4**, 215–32.

(*Manuscript received* 15 *October* 1957)

THE EFFECT OF A WARNING SIGNAL ON UNRESTRICTED AVOIDANCE BEHAVIOUR*

By J. D. KEEHN

American University of Beirut, Lebanon

In an avoidance situation where interval responses served to postpone the onset of the next 'trial' animals were trained with and without a warning signal. Those animals which received the warning signal were free to postpone the onset of this signal as well as shock. It was found, however, that these animals come to respond more and more in the presence of the signal instead of before it. That is, they behaved so as to postpone the shock but not the warning signal. This finding is interpreted as showing that, contradictory to the opinion of many psychologists, the warning signal does not acquire secondary aversive properties. Instead it is argued that this signal has the properties of a discriminative stimulus in the sense that it marks the occasions on which the appropriate avoidance response is reinforced. The relevance of this point of view to paradoxical statements by Hull and Pavlov is pointed out.

I. Introduction

A. *Theoretical status of avoidance behaviour*

In the recent history of learning theory the phenomenon of so-called avoidance conditioning has occupied a prominent place. Findings from experiments on this kind of behaviour have been related, for instance, to the contiguity-reinforcement controversy via studies on the relative efficiency of the classical and instrumental conditioning procedures (Brogden, Lipman & Culler, 1938; Sheffield, 1948). They pose problems for theorists attempting to maintain a consistent monistic drive reduction theory (Miller, 1951*b*; Mowrer, 1951). They provide the framework within which the merits and demerits of a two-factor theory of learning have been debated (Mowrer, 1951; Bugelski, 1956; Kendler, 1951), and they pertain to the role of fear in motivating avoidance behaviour (Solomon & Wynne, 1954; Schoenfeld, 1950; Woodworth & Schlosberg, 1954; Mowrer, 1947; Skinner, 1953). It is with this last point that the present paper is particularly concerned.

In order to maintain a drive reduction theory of reinforcement Mowrer (1940) introduced the notion of fear conditioning into the avoidance paradigm. Miller (1948) subsequently provided the reference experiment, and the view that avoidance behaviour came about by, first, the conditioning of fear to apparatus and other stimuli and, secondly, the reduction of this fear by the avoidance response gained ground. Mowrer (1951) and Miller (1951*a*) disagreed on the manner in which fear became conditioned, but the matter seemed to be settling itself in Mowrer's favour (Mowrer & Solomon, 1954; Mowrer & Aiken, 1954). However, a new development that has come about in recent years has set the phenomenon of avoidance behaviour in an entirely new light. For many years now it has been known that, in addition to 'conditioned' or signalled responses in avoidance learning experiments, there also occurred interval, or non-signalled responses of like kind (Mowrer & Lamoreaux, 1942; May, 1948). Moreover, aside from his main

---

* Part of this study was carried out while the writer was on leave as a Ford Fellow at the University of Illinois during the year 1956–7. Acknowledgement is made to the Ford Foundation and to Prof. O. H. Mowrer whose assistance and encouragement made the study possible.

research, May noticed that if these interval responses were allowed to postpone the next trial they tended to become persistent. May's observations were utilized by Bugelski & Coyer (1950) in devising a technique to demonstrate what they called 'temporal conditioning', but the main advances along these lines have come from a procedure developed by Sidman (1953).

B. *Interval responses made functional*

Basically, Sidman's technique was to give animals periodic brief shocks with no prior 'warning' stimulus unless they made some pre-determined response, such as pressing a lever, in the intervals. Each response functioned to postpone the next shock and it was possible for the animals, by responding above some minimum rate, to avoid shock altogether. Quite successful avoidance behaviour was 'conditioned' by this method. Sidman interpreted his findings in terms of a theory of avoidance behaviour advanced by Schoenfeld (1950). According to Schoenfeld, avoidance responding comes about in two ways. In the first place, all behaviour other than the successful avoiding response is, sooner or later, followed by shock and thereby punished: thus the particular avoidance response becomes dominant not through any direct reinforcement to itself but through the relative depression of all other activity. In the second place, the correct response is secondarily rewarded in that it removes the proprioceptive stimuli left by other, punished, responses.

An alternative account of these findings was offered by Mowrer & Keehn (1958) who assumed that fear became conditioned to some point on the decaying stimulus trace left by the avoidance response (which was also the escape from shock response) so that, at a time after a response had been made, fear would mount and trigger off the next response and so on. This point of view required that their subjects, white rats, be able to discriminate an interval of about 20 sec., and such has been shown to be the case (Sidman, 1956).

Despite their differences, both these interpretations of avoidance behaviour attribute aversive properties, either directly or indirectly, to stimuli present at the onset of shock. However, it has been shown by Sidman (1955) and by Sidman & Boren (1957) that when a warning signal is introduced into the postponing procedure such that responses can postpone both shock and stimulus the majority of responses tend to occur during, rather than prior to, the onset of the warning stimulus. In other words, if these findings are reliable, animals do not postpone a warning signal when they are given the opportunity but, instead, tend to suppress responses in its absence. Yet if *aversive* properties are attributed to a stimulus preceding shock induction we should expect animals to respond so as to postpone the onset of the signal just as they react to delay the onset of shock.

Unfortunately, experimental findings to date are not conclusive on the matter for, in every case, before a warning signal has been introduced considerable training without it has been given and a good deal of presignal responding has continued to occur. The present study was designed to test the different manners of responding by animals trained wholly with or wholly without a warning signal.

II. METHOD

Experiment 1. The apparatus was an ordinary activity wheel of diameter 14 in. and width 5 in. which was wired for shock. The wheel could turn in one direction only. Electrical timer and relay circuits housed in sound-absorbing boxes lined with fibre-glass served to energize the grid floor of the wheel at 20 sec. intervals. A prescribed turn of the wheel, however, automatically reset the shock timer. Hence shocks could occur every 20 sec. (S–S interval) and every response served to postpone the next shock by 20 sec. (R–S interval). The same running response that reset the timer terminated the shock (175 V.; 250,000 Ω limiting resistor) when it was not avoided.

Response records were made by means of a motor-driven pen moving at constant speed across a drum geared directly to the activity wheel. Movement of the wheel turned the drum relative to the pen thus automatically producing a cumulative response curve. Revolutions were recorded separately.

Four pairs of male albino rats from the University of Illinois colony were trained individually when they were between 140 and 160 days old. Two pairs were given three 2½ hr. training sessions on alternate days. The others had eight daily sessions of 1 hr. duration. All animals were given two adaptation sessions of the same duration as those in which they were to be trained. For one animal (S) in each pair a tone stimulus came on 4 sec. before a shock might be given and lasted until the animal responded either to it alone or in combination with shock. Interval responses postponed both shock and tone for these animals. No exteroceptive stimulus was used with the second animal (NS) in each pair.

Experiment 2. While Exp. 1 was designed to yield data on the overall activity of the subjects and the form of their cumulative response curves, no information was gathered on the effect of training on inter-response times. Although these times can be inferred from the cumulative response curves this can only be done with difficulty and with a high degree of inaccuracy. It was to remedy this deficiency that a second experiment was conducted.

Three pairs of male albino rats from the American University of Beirut Medical School colony were trained after the manner described above, but on this occasion the procedure was operated by hand instead of electronically.* Training sessions were 1 hr. long and continued for eight successive days for each animal. Although a different activity wheel was used in this case its dimensions were the same as that used in Exp. 1.

Record was kept of the periods spent in resting between runs, the total number of revolutions that each animal turned the wheel in each session, and the number of shocks it received.

III. RESULTS

Experiment 1. In all pairs of animals the S animal consistently responded at a lower rate than its partner. The response curves shown in Figs. 1 and 2 are typical and are the complete records made by one pair of animals trained with 2½ hr. sessions. The horizontal lines marked A2 at the left of each figure show the amount that the respective animals ran on the second adaptation day. As these curves show, hardly any running occurred on this day, animal 21 turned the wheel only three times and animal 61 not at all. Other writers (Kagan & Berkun, 1954; Seward & Pereboom, 1955) have reported much higher spontaneous running rates of albino rats, and our own experience with young Sprague-Dawley males suggests strain differences for this kind of behaviour. The curves marked T1, T2 and T3 represent the amount the animals ran on the respective training days. As each record was made on a circular roll of paper the curves within each set are continuous, each line representing one complete or partial revolution of the paper.

On the three successive training days the S animal (no. 21) turned the wheel 408, 378 and 356 revolutions respectively. Over the same period the number of shocks received, marked by lateral spurs on the graphs, dropped from 33 in the first session to 5 in the second and 6 in the third. Similar data on its NS partner (no. 61) are 586, 650 and 562 revolutions with 96, 5 and 9 shocks respectively. Combining the responses of all the animals over all the training sessions the NS group totalled 5655 revolutions of the wheel as against 3628 by their S partners. (One animal on the hourly training schedule died after 5 days of training, so the data from it and its partner are not included.) Similarly, the NS animals received a grand total of 234 shocks while their partners were shocked a total of 82 times.

---

* This experiment was made possible by a grant from the Rockefeller Foundation to the Arts and Sciences Division of the American University of Beirut. Grateful acknowledgement is made to this foundation and to Miss U. Sabbagh who collected the data.

Although the S animals responded at a lower rate than their partners these animals still emitted many responses before the onset of the signal, thus postponing it as well as the shock. This is clearly shown in Fig. 3, which was constructed by plotting the cumu-

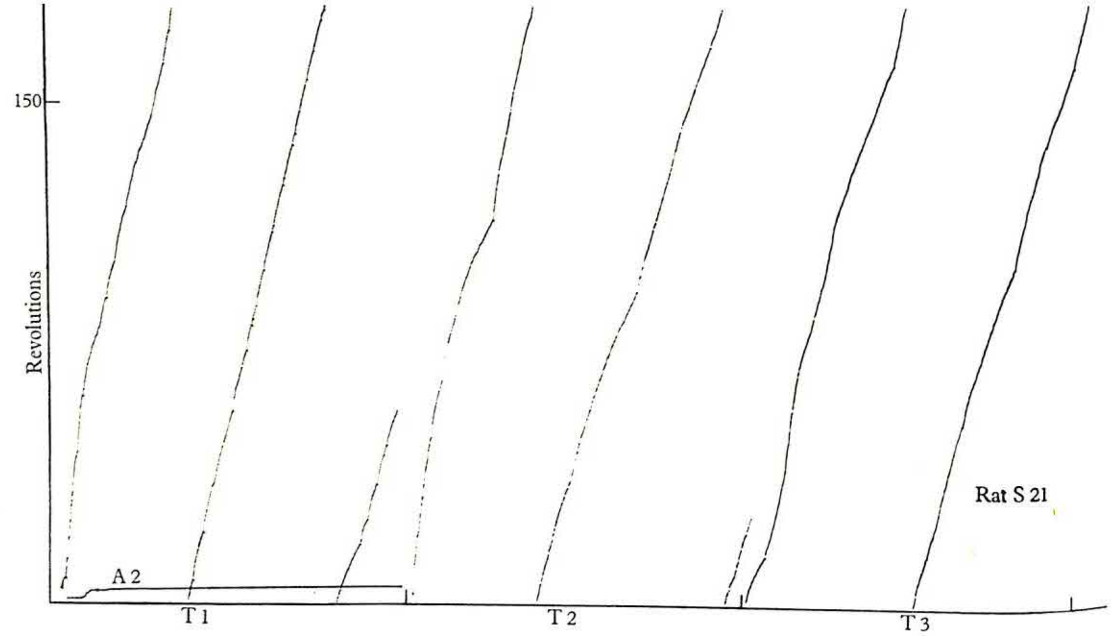


Fig. 1. Cumulative running (avoidance) behaviour on the second adaptation day, A 2, and training sessions T 1, T 2 and T 3. See text for details.

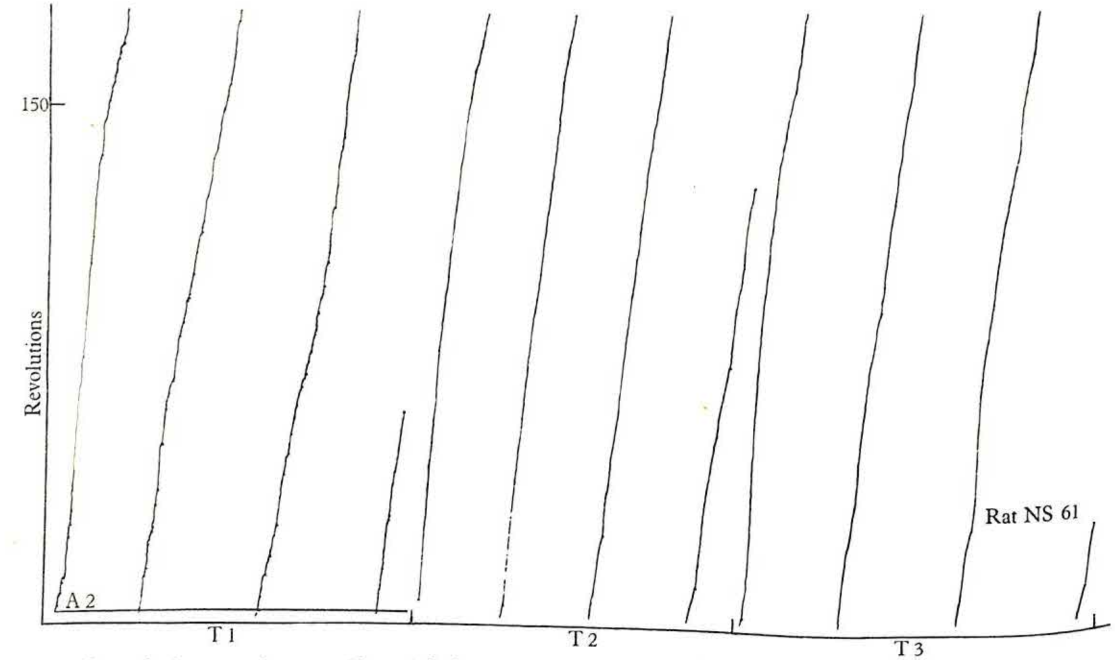


Fig. 2. Cumulative running (avoidance) behaviour on the second adaptation day, A 2, and training sessions T 1, T 2 and T 3. See text for details.

lative number of non-shocked responses made to the tone against time on successive training days for the S animal which completed the 8-hourly training sessions. The 'optimum' curve is based upon responses made only with the tone on, i.e. about one every 20 sec. Extra, pre-tone, responses by postponing the tone as well as the shock necessarily depress the number of tones sounded during the session and hence the number of responses made to it. Thus if the animal had responded only in the presence of the tone the cumulative response curves would have coincided with the 'optimum' response curve. From Fig. 3 it can be seen that the number of responses made to the tone approaches the optimum value as training increases, although the animal's behaviour was more efficient on the penultimate than on the final training day.

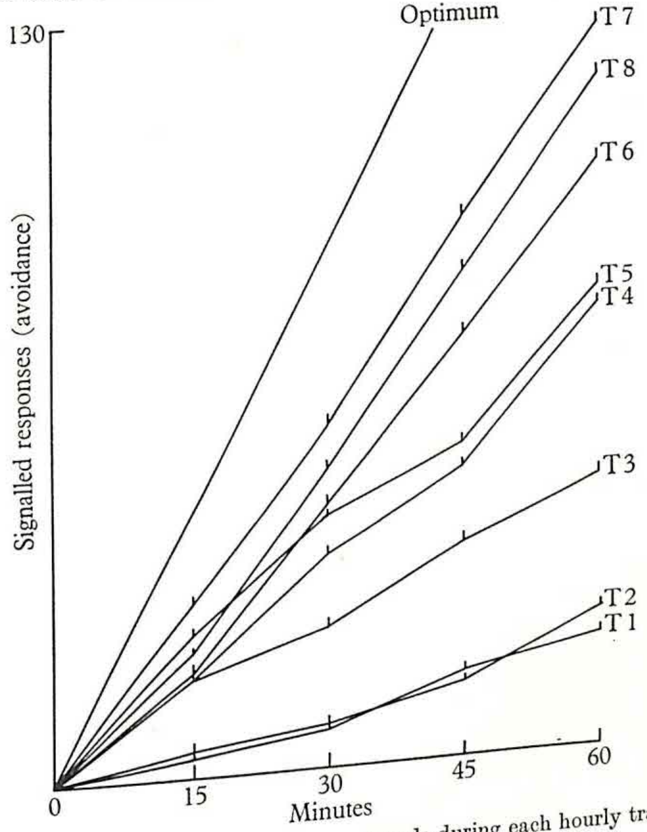


Fig. 3. Cumulative number of signalled responses made during each hourly training session for the sessions T1 to T8. See text for details.

Experiment 2. Table 1 shows the principal data yielded by Exp. 2. Once again, the combined S animals turned the wheel fewer times than the NS group (3563 versus 3938), although in this instance one NS animal, no. 3, ran less than its partner. However, this animal was abnormal among those used in our experiments in that it did not move at all during its adaptation trials, ran hardly at all on its first training trial, and received considerably more shocks than any other animal.

Of more particular interest is the effect of the signal on the intervals between responses. Fig. 3 shows the increase in the number of times within and between sessions that one S animal waited between 16 and 20 sec. before responding. Table 1 contains similar information, for between sessions only, for all those pairs of animals used in the present experiment. It is clear that long delays between responses are considerably more frequent in the case of the S animals. Even the S3 animal which ran more than its partner rested for longer periods; 717 against 150. The S group responded to the tone 1847 times in all, whereas delays between 16 and 20 sec. occurred only 261 times throughout the eight training sessions of the NS animals. Despite these longer rest periods the S animals received fewer shocks than their partners; 206 as against 446.

Table 1

Animals		Training days								
		T1	T2	T3	T4	T5	T6	T7	T8	Total
S1	(A)	26	58	73	97	97	115	106	89	661
	(B)	60	10	16	4	5	12	11	15	133
	(C)	106	116	98	115	107	60	102	118	822
S2	(A)	46	58	62	78	65	42	47	71	469
	(B)	22	3	2	1	4	8	0	1	41
	(C)	125	168	91	82	130	155	160	108	1019
S3	(A)	90	106	51	122	114	68	111	55	717
	(B)	21	4	1	2	2	1	1	0	32
	(C)	337	81	252	77	169	286	192	328	1722
NS1	(A)	16	4	2	3	3	0	8	3	39
	(B)	31	6	4	2	2	0	1	4	50
	(C)	139	206	130	124	141	139	130	170	1179
NS2	(A)	2	11	13	9	11	12	6	8	72
	(B)	98	50	9	4	10	4	4	3	182
	(C)	386	248	126	138	235	162	325	258	1878
NS3	(A)	0	39	16	11	21	9	30	24	150
	(B)	121	20	23	9	10	10	9	12	214
	(C)	25	73	129	181	139	73	139	122	881

Row (A) indicates for each animal the number of responses made after a delay of 16–20 sec. after the preceding response. In the case of the S animals these are non-shocked responses occurring in the presence of the tone. Delays of greater than 20 sec. are always shocked, and the number of shocks received are recorded in the respective (B) rows. Row (C) in each case shows the number of revolutions that each animal turned the activity wheel on the training days T1 to T8.

IV. Discussion

(a) *The effect of the warning signal*

The first point of note is the highly effective avoidance behaviour generated by our technique. For instance, in the very first $2\frac{1}{2}$ hr. session an animal might take 40 or 50 shocks (often considerably fewer) out of a possible 450, and after only three such sessions avoid all but five or six shocks. This is considerably more efficient behaviour than that obtained with conventional avoidance training procedures, or even with the technique used by Sidman. Furthermore, these results are obtained with what amounts to a *massed practice* procedure, an approach which is usually thought to decrease the rate of learning or responding.

Of more immediate concern herein is the effect that the warning stimulus has on the rate of responding. Fewer responses are emitted by animals which are 'warned' by the onset of a tone that a shock is due in 4 sec. unless it behaves appropriately than by animals given no such warning. Furthermore, this comes about through an increase in the number of long latency responses made by the S animals (see Table 1). That is, these animals instead of behaving so as to postpone the signal as well as the shock come to respond more and more only in the presence of the tone. Whether or not responding would come to concentrate *only* in the period during which the tone sounds cannot be said from our present data. The evidence shown in Fig. 3 suggests that this might be the case, but the data in Table 1 indicate that non-signalled responses are difficult to eliminate entirely, a finding which supports those of Sidman (1955) and Sidman & Boren (1957) who found that non-signalled responses were not eliminated even after many hours of training.

Thus it is evident that animals do not behave so as to postpone the onset of an allegedly aversive stimulus when they are given the opportunity to do so. Instead they tend more and more, up to a point, to concentrate their responses in the period when this stimulus is operating. A closer examination of the properties of such a stimulus thus becomes imperative and to this end an examination of what might be called *observing responses*, Wyckoff (1952), is helpful.

(*b*) *Observing responses*

If a formerly neutral stimulus consistently precedes some noxious event, animals come to behave so as to eliminate this stimulus if, thereby, the noxious event is avoided. Clearly such behaviour is biologically advantageous. As Pavlov (1927) says, 'The strong carnivorous animal preys on weaker animals, and these if they waited to defend themselves until the teeth of the foe were in their flesh would speedily be exterminated. The case takes on a different aspect when the defense reflex is called into play by the sights and sounds of the enemy's approach. Then the prey has a chance to save itself by hiding or by flight' (p. 14). Behaviour of this kind is interpreted by Pavlov in classical conditioning terms, whereby the 'defense reflex' becomes conditioned to the 'sight and sounds' of the predator. This explanation is unsatisfactory, however, for unless an animal experience one or more attacks there would be no opportunity for conditioning to occur. If, as Pavlov believes, conditioning serves to improve an animal's chances of survival it can only function for those animals which have managed to survive by some other means. Moreover, as the conditioned properties of the stimulus extinguish when they are not reinforced (owing to the avoidance of attack in this case) the animal would quickly be exposed to further possibly lethal experiences. Furthermore, specific laboratory studies of avoidance behaviour (Mowrer & Lamoreaux, 1946; Mowrer & Keehn, 1958; Keehn, 1959), have demonstrated the inadequacies of a simple Pavlovian interpretation in this case, but in every instance alternative explanations have relied upon attributing aversive properties to the 'warning' stimulus.

This being so, we should expect animals to behave so as to avoid this stimulus also. Not only, as has been shown above, do animals not do this but also it would be biologically disadvantageous if they did. If there is survival value in reacting to the 'sights and sounds of the enemy's approach', it is also useful to be on the lookout for these sights and sounds. The animal which remains oblivious to the approach of a predator is soon a dead animal. From the standpoint of biological utility, then, it is (up to a point) better for the animal to indulge in activity which would, to all intents and purposes, turn on a secondary aversive stimulus, or fear arouser. Yet the concensus of psychological opinion is that activity becomes such as to turn off, not turn on, aversive stimuli. The reason for this is not hard to find. In laboratory experiments on avoidance behaviour stimuli have always been used which could not fail to impinge upon the animal whatever it might be doing. Thus a tone is sounded or the gross illumination changed and no opportunity given the animal not to 'notice' these events. When such opportunity is given, by using a small light behind a screen, for instance, rats do learn to 'look for' the stimulus, although with some difficulty.

Properties instead of, or in addition to, fear arousal must therefore be attributed to signals preceding noxious events. If the properties of a *discriminative stimulus* (Skinner, 1938) are given to such a signal, a solution to the paradox is forthcoming. Imagine a young naïve ruminant and its parent quietly grazing in their natural habitat. A predator appears on the scene. If the offspring stays unperturbed it will be pushed and prodded and nipped by the parent until it finally runs for shelter. At irregular intervals the youngster suffers these indignities until finally it comes to respond *only* in the predator's presence as it is only on such occasions that it is punished by its parent. Ultimately, then, the young animal flees at the appearance of its foe and avoids the pain inflicted by the parent. Fleeing is not conditioned to the 'sights and sounds' of the enemy, nor need the foe be feared. These stimuli possess only the properties of *discriminative stimuli* in that they set the occasion for the reinforcement of the appropriate response. As punishment by the parent occurs only in the (subjective) absence of these stimuli it is reasonable to expect the animal to 'look for' the predator and to run only in its presence.

(c) *Properties of the warning signal in laboratory studies*

Earlier in this paper we demonstrated that, under the proper conditions, animals come to respond more in the presence of a stimulus that had preceded shock than before it, and noted that this contradicted the notion that this stimulus should possess aversive properties. Secondly, we noted that by attributing discriminative instead of aversive properties to such a stimulus we could account for the natural observation that animals behave so as to be on the look-out for the approach of a predator. The question now arises as to whether we need attribute *only* discriminative properties to the signal used in laboratory studies of avoidance (postponing) learning.

According to Sidman (1957), this is all that is necessary. Says this writer: 'Sidman & Boren (1957) have suggested that the discriminated avoidance situation may be considered as a "multiple schedule" in which one avoidance contingency prevails in the presence of the warning stimulus and another in its absence. (This) viewpoint places the explanatory stress upon the operant consequences of the behavior involved. The warning stimulus, rather than being simply a vehicle for the Pavlovian conditioning of drives or aversive properties, becomes a discriminative stimulus that enters into an operant contingency.' However if, as this passage suggests, avoidance activity is analogous to the food-producing behaviour of hungry animals in the usual kind of Skinnerian experiment some *continuous* motivational source comparable to that of hunger must be postulated. At our present state of understanding, *fear* is the most likely candidate for this role and the problem becomes one of defining the fear-arousing stimulus. From the kinds of experiments usually performed to date only two alternatives are possible. Either fear is aroused by the specific warning signal or it is conditioned to the whole situational stimulus compound. Both the experimental findings reported above and our analysis of 'observing behaviour' suggest that the warning signal itself is not aversive and so we are left with the latter alternative. That is, in both the natural and the experimental settings, animals acquire, either by parental punishment or periodic shocks, a certain level of 'anxiety' and react so as to reduce this anxiety. In the case where a specific stimulus precedes the noxious event this stimulus comes to distinguish between occasions when responses will be reinforced and occasions when they will not. Only in the event that 'signalled' responses receive a greater amount of reinforcement than non-signalled ones (by postponing punishment for a longer time) will the former be expected to outweigh the latter and such has been shown to be the case by Sidman (1957).

V. CONCLUSIONS

Several interesting possibilities develop out of the kind of analysis given above. In the first place, it bears upon the contradictory properties given by Hull (1952) to a formerly neutral stimulus which precedes reinforcement. When referring to experiments like that of Cowles (1937), where chimpanzees are trained to work for coloured tokens, such a stimulus is given the status of a secondary reward. On the other hand, when a stimulus precedes electric shock (Miller, 1948) it is said to come to initiate a secondary drive. Many writers, including Mowrer (1952), Miller (1951*b*), Brown (1953) and Hilgard (1956, p. 129), have pointed to the contradictory nature of these two postulates and a variety of resolutions of the contradictions have been proposed. If the pre-shock signal in the Miller experiment is seen as a discriminative stimulus Hull's difficulty disappears, especially if, as Dinsmoor (1950) has argued, discriminative stimuli and secondary reinforcers are one and the same thing.

In the second place, it is evident that the conventional avoidance procedure is but a special case of the more general kind of postponing behaviour generated by Sidman's (1953) technique. Thus, while we began this paper with reference to 'interval' responses occurring between 'trials' in the usual avoidance learning experiment (May, 1948), it is now possible to see that the so-called 'true' avoidance response is simply an interval response (between shocks) that coincides with a pre-shock signal. This signal serves not to cause the response but to mark the occasions on which responses will be reinforced, and the reinforcement comes about not through the avoidance of shock but through its postponement—from trial to trial.

As to the precise nature of the reinforcement in this case it is not easy to say. We have preferred to talk in terms of something like 'anxiety reduction', but the problem remains as to how anxiety is re-instated. It may be, as Mowrer & Keehn (1958) say, that fear rises as the proprioceptive traces of previous responses decay, but the data presented herein cast doubt upon such an explanation, at least in so far as 'external' stimuli are concerned. Alternatively, it is possible that anxiety is reduced every time a correct response is made until, finally, no response is made and the animal is shocked and the anxiety, or fear, level reinstated. What is important is the demonstration of an entirely new approach to the mechanics of motivation and the operation of reinforcement. It has always been difficult, for instance, for psychologists to maintain that human adults eat *because* they are hungry; we might conceive of eating as serving to *postpone* hunger. In the cases where feeding has occurred on an irregular schedule such as might be expected to generate a high level of 'food-anxiety' it is not difficult to see why eating, sucking, or other related behaviours might occur at an unusually high rate without recourse to Freudian mechanisms (see Hunt, 1941).

REFERENCES

BROGDEN, W. J., LIPMAN, E. A. & CULLER, E. (1938). The role of incentive in conditioning and learning. *Amer. J. Psychol.* **51**, 109–17.

BROWN, J. S. (1953). Problems presented by the concept of acquired drives. In *Current Theory and Research in Motivation.* Lincoln: University of Nebraska Press.

BUGELSKI, B. R. (1956). *The Psychology of Learning.* New York: Holt.

BUGELSKI, B. R. & COYER, R. A. (1950). Temporal conditioning *vs.* anxiety reduction in avoidance learning. *Amer. Psychol.* **5**, 264–5. (Abstract.)

COWLES, J. T. (1937). Food-tokens as incentives for learning by chimpanzees. *Comp. Psychol. Monogr.* **14**, no. 5.

DINSMOOR, J. A. (1950). A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. Exp. Psychol.* **40**, 458–72.

HILGARD, E. R. (1956). *Theories of Learning* (2nd edition). New York: Appleton-Century-Crofts.

HULL, C. L. (1952). *A Behavior System.* New Haven: Yale University Press.

HUNT, J. McV. (1941). The effects of infant feeding-frustration upon adult hoarding behavior. *J. Abnorm. Soc. Psychol.* **36**, 338–60.

KAGAN, J. & BERKUN, M. (1954). The reward value of running activity. *J. Comp. Physiol. Psychol.* **47**, 108.

KEEHN, J. D. (1959). On the non-classical nature of avoidance behavior. *Amer. J. Psychol.* (in the Press).

KENDLER, H. H. (1951). Reflections and confessions of a reinforcement theorist. *Psychol. Rev.* **58**, 368–74.

MAY, M. A. (1948). Experimentally acquired drives. *J. Exp. Psychol.* **38**, 66–77.

MILLER, N. E. (1948). Studies of fear as an acquirable drive. I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. Exp. Psychol.* **38**, 89–101.

MILLER, N. E. (1951*a*). Comments on multiple-process conceptions of learning. *Psychol. Rev.* **58**, 375–81.

MILLER, N. E. (1951*b*). Learnable drives and rewards. In Stevens, S. S. (ed.). *Handbook of Experimental Psychology.* New York: Wiley.

MOWRER, O. H. (1940). Anxiety reduction and learning. *J. Exp. Psychol.* **27**, 497–516.

MOWRER, O. H. (1947). On the dual nature of learning—a reinterpretation of 'conditioning' and 'problem solving'. *Harvard Educ. Rev.* **17**, 102–48.

MOWRER, O. H. (1951). Two factor learning theory: summary and comment. *Psychol. Rev.* **58**, 350–4.

MOWRER, O. H. (1952). Motivation. In Stone, C. P. & Taylor, D. W. (eds.). *Annual Review of Psychology.* Stanford: Annual Reviews, Inc.

MOWRER, O. H. & AIKEN, E. G. (1954). Contiguity *vs.* drive-reduction in conditioned fear: temporal variations in conditioned and unconditioned stimulus. *Amer. J. Psychol.* **67**, 26–38.

MOWRER, O. H. & KEEHN, J. D. (1958). How are inter-trial 'avoidance' responses reinforced? *Psychol. Rev.* **65**, 209–21.

MOWRER, O. H. & LAMOREAUX, R. R. (1942). Avoidance conditioning and signal duration: a study of secondary motivation and reward. *Psychol. Monogr.* **54**, no. 5.

MOWRER, O. H. & LAMOREAUX, R. R. (1946). Fear as an intervening variable in avoidance conditioning. *J. Comp. Physiol. Psychol.* **39**, 29–50.

MOWRER, O. H. & SOLOMON, L. N. (1954). Contiguity *vs.* drive-reduction in conditioned fear: the proximity and abruptness of drive-reduction. *Amer. J. Psychol.* **67**, 15–25.

PAVLOV, I. P. (1927). *Conditioned Reflexes.* London: Oxford University Press.

SCHOENFELD, W. N. (1950). An experimental approach to anxiety, escape and avoidance behavior. In Hoch, P. H. & Zubin, J. (eds.). *Anxiety.* New York: Grune and Stratton.

SEWARD, J. P. & PEREBOOM, A. C. (1955). Does the activity wheel measure goal striving? *J. Comp. Physiol. Psychol.* **48**, 272–7.

SHEFFIELD, F. D. (1948). Avoidance training and the contiguity principle. *J. Comp. Physiol. Psychol.* **41**. 165–77.

SIDMAN, M. (1953). Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science*, **118**, 157–8.

SIDMAN, M. (1955). Some properties of the warning signal in avoidance behavior. *J. Comp. Physiol. Psychol.* 48, 444–50.

SIDMAN, M. (1956). Time discrimination and behavioral interactions in a free operant situation. *J. Comp. Physiol. Psychol.* **49**, 469–73.

SIDMAN, M. & BOREN, J. J. (1957). A comparison of two types of warning stimulus in an avoidance situation. *J. Comp. Physiol. Psychol.* **50**, 282–7.

SIDMAN, M. (1957). Conditioned reinforcing and aversive stimuli in an avoidance situation. *Trans. N.Y. Acad. Sci.* Ser. II, **19**, 534–44.

SKINNER, B. F. (1938). *The Behavior of Organisms.* New York: Appleton-Century-Crofts.

SKINNER, B. F. (1953). *Science and Human Behavior.* New York: MacMillan.

SOLOMON, R. L. & WYNNE, L. C. (1954). Traumatic avoidance learning: The principles of anxiety conservation and partial irreversibility. *Psychol. Rev.* **61**, 353–85.

WOODWORTH, R. S. & SCHLOSBERG, H. (1954). *Experimental Psychology.* London: Methuen.

WYCKOFF, L. B. (1952). The role of observing responses in discrimination learning. Part I. *Psychol. Rev.* **59**, 431–42.

(*Manuscript received* 9 *January* 1958)

EFFECTS OF AGE AND EXPERIENCE ON THE ESTABLISHMENT OF INTERNAL INHIBITION IN FINCHES

By M. A. VINCE

Cambridge Psychological Laboratory, and Ornithological Field Station, Madingley

An experiment was carried out to test the hypothesis that internal inhibition, learning not to respond in an un-reinforced situation, is more stable in adult than in young birds and also more stable in birds reared under more varied conditions than in those reared under restricted conditions. Canaries and greenfinches were given the task of differentiating between a dish containing food and covered with a white lid, and an empty dish covered with a black lid. In the earlier negative trials, juvenile canaries were superior to adults in their capacity to ignore the black lid, and juvenile hand-reared greenfinches were superior to adult wild greenfinches; but in the later trials the two adult groups were superior to these two juvenile groups. A group of juvenile greenfinches reared in an outdoor aviary were superior to all other groups. These results are discussed and it is suggested that a more detailed hypothesis may be formulated: that internal inhibition is weak in very young birds, develops as a result of age and also as a result of experience during the juvenile stage and, later again, weakens slightly.

I. INTRODUCTION

In a situation where they could obtain food by pulling up a string, canaries and greenfinches gave results which varied according to the age at which they were tested (Vince, 1958). In each half-hour test period, the time spent responding to string and/or food tended to diminish in all birds. In young birds, however, it returned to its initial high level at the beginning of the next day's test period, while in adults the diminution tended to continue from day to day, and in some individuals the response ceased altogether. Young birds reared in an outdoor aviary gave results more like those obtained from the adults than those from young birds reared under the more restricted conditions indoors. It was suggested that this diminution of the response, occurring as it did only in birds which failed to obtain the bait, was an example of Pavlov's extinction of conditioned reflexes; hence the results gave rise to the hypothesis (i) that internal inhibition is unstable in young birds, and (ii) that the development of this process from unstable to stable might depend on experience rather than age alone. It was decided to test this hypothesis in a different (differentiation) experimental situation, but one which also, according to Pavlov (1927), depended on the establishment of internal inhibition. Most of the birds tested were used in the previous experiment, although in the present experiment they were a few weeks older.

II. EXPERIMENTAL SITUATION

The birds were tested indoors in flights measuring approximately 3 ft. × 3 ft. × 3 ft. They were allowed to feed throughout the whole period of testing, but the food used as bait (chickweed) was not given to them at any other time.

1. *The initial training.* Before testing began the birds were trained to feed by removing the white cardboard lid from a dish containing food. Some chickweed was placed in a blue dish $3\frac{1}{2}$ in. in diameter at the top. This was placed on the floor of the flight and the time within which the bird went to feed was recorded. When the bird fed readily from the dish it was presented in the usual way, but covered with a white cardboard lid. Some birds continued to alight on the rim of the dish, and soon began to feed by first removing the lid, but some required a fairly long period of training before they were successful. Some could not be trained. A successful response is shown in Fig. 1. The numbers trained to remove

the lid and the number which failed are given in Table 1, together with the average latent period for each group before testing began, and the average number of trials required by each group before the birds removed the lid immediately on its being placed in the flight.

2. *Species, age and previous experience of birds tested.* Birds which succeeded in the initial training were subsequently tested in five groups as shown in Table 1. The birds, of which eleven were domestic canaries (*Serinus*) of the Border strain and fourteen were greenfinches (*Chloris chloris*) varied in respect to age and/or early experience. The rearing and subsequent experience of the juvenile canaries was similar to that of the adult canaries; they varied in age alone and were tested in two groups. Group (i) consisted of five adults, and group (ii) of six juveniles aged between 11 and 29 weeks at the beginning of the test trials. The greenfinches were tested in three groups: (i) five wild-reared birds which had been trapped and kept in aviaries; (ii) five juveniles which had been taken as wild nestlings were then reared by hand, kept indoors and tested from 14½ to 16 weeks of age; and (iii) four juveniles reared by wild trapped adults in an out-door aviary, where they remained until required for testing; these were aged between 11½ and 13 weeks at the beginning of the test period. The two groups of juvenile greenfinches were therefore of comparable age, but varied in early experience; the wild and aviary-reared greenfinches, on the other hand, varied in age, but were reared under conditions which were roughly comparable. In addition to the birds already mentioned, four canaries (two adult and two juvenile) were tested, but scored only for success or failure. Data obtained from these birds have been included in the results where latencies are not considered.

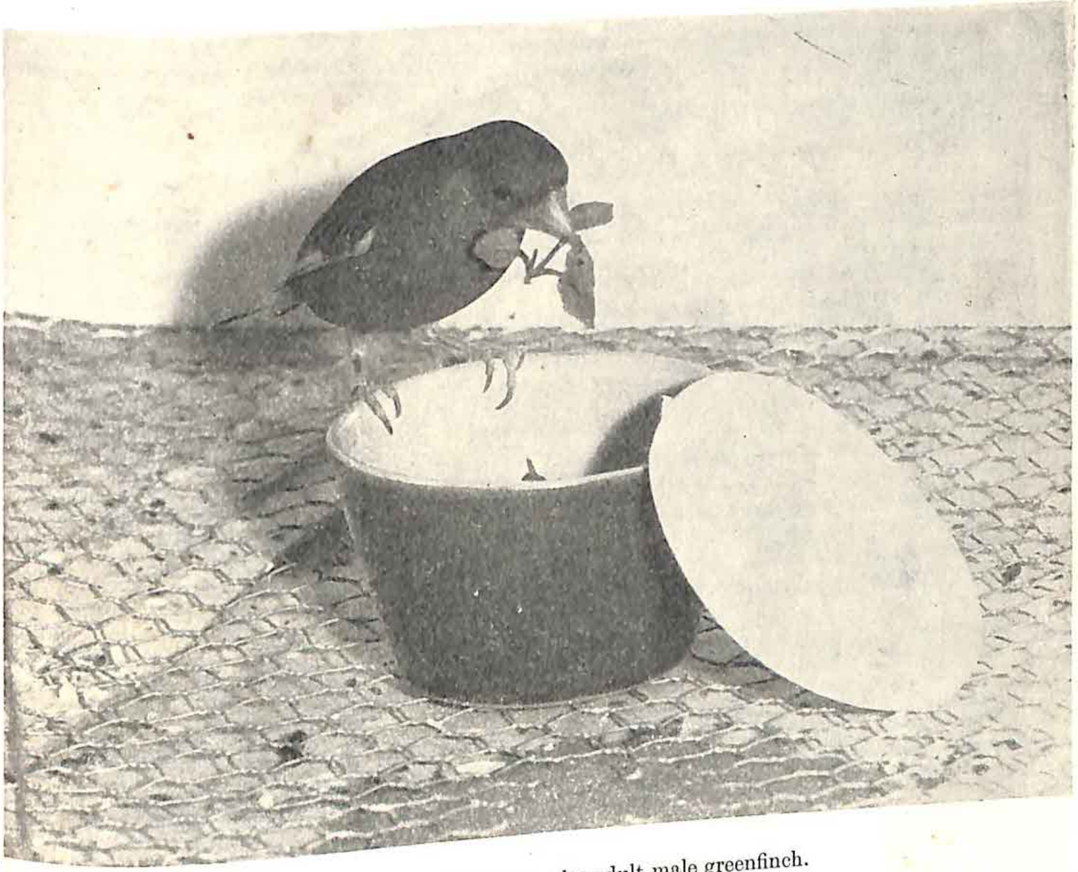


Fig. 1. Successful response by adult male greenfinch.

3. *The test period.* The birds were given three test trials a day, the interval between trials being about one hour. Each trial consisted of the presentation, for five minutes, of the dish covered with a black or white lid. The test period began with the first presentation of the empty dish covered with the black lid. In most cases, the lid-removing response was generalized to this, although after an unusually long latent period on the first occasion. The order of presentation of the black and white lids was

randomized as far as possible, but an approximately equal number of positive and negative trials was given and on no day were all three trials either positive or negative. Testing was continued until each bird had had at least twenty negative trials, and by this time most birds had at least begun to ignore the dish with the black lid.

Three birds failed to complete the testing, as shown in Table 1. In each greenfinch group one bird failed to respond at all to the black lid. The data obtained from these three birds have been omitted from the discussion below.

Table 1

	Canaries		Greenfinches		
	Adult	Juvenile	Adult	Juvenile hand-reared	Juvenile aviary-reared
Training period					
1. No. given initial training in removing white lid from dish containing chickweed	11 (two scored only for success or failure)	9 (two scored only for success or failure)	8	5	5
2. No. which failed to learn to remove lid	4	1	3	0	1
3. Average no. of trials required for training (successful birds only)	15·4	9·6	44·6	4·6	27·3
4. Mean latent period in last three training trials (time within which white lid or dish was touched)	17 sec. (five birds only)	20 sec. (six birds only)	9 sec.	18 sec.	7 sec.
Test period					
5. No. of birds tested with white (re-inforced) and black (un-reinforced) lids	7	8	5	5	4
6. Age of juveniles at time of first negative trial	—	11–29 weeks	—	14–16 weeks	11–13 weeks
7. No. which did not remove black lid in five negative trials and were abandoned	0	0	1	1	1
8. No. included in results	7	8	4	4	3
9. No. which did not respond to black lid on first negative trial	1	1	1	4	1
10. Average latency of response to black lid on first negative trial	101 sec. (five birds only)	154 sec. (six birds only)	275 sec.	360 sec.	311 sec.
11. Total no. of positive trials in which there was no response to white lid	6	12	0	2	0
12. Average latency of response in twenty positive trials	25 sec. (or 4·6 if one bird omitted)	21 sec.	8·5 sec.	24 sec.	16 sec.

III. RESULTS

The records of one adult and one juvenile canary are given in Fig. 2A and B, showing the latent period (time elapsing before dish or lid was touched by the birds) for successive positive and negative trials. A correct positive response is indicated by a short latency and a correct negative response by a long latency.

The significant scores for the purposes of this investigation are the latent period in the negative, un-reinforced trials, and the number of correct negative responses (no response to black lid in five minutes) scored by each bird. They show how rapidly and how consistently the birds came to ignore the black lid indicating an empty dish. The latency of response in positive, reinforced trials, or alternatively the number of incorrect positive trials where the white lid was ignored for five minutes, will also have to be considered; the stability of a differentiation clearly depends not only on a capacity to ignore the negative stimulus, but also on continued prompt response to the positive one.

It is possible to look at the results in a number of different ways. If we consider the average latency over the twenty negative trials (Table 2), the greenfinches were superior to the canaries and the aviary-reared were superior to the other greenfinches, but there were no appreciable differences between the scores of adult and juvenile canaries, or between adult wild and juvenile hand-reared greenfinches. If we consider the percentage of correct negative trials (no response to the black lid in five minutes, see also Table 2), the results are similar, although here the adults were somewhat superior to the juveniles, suggesting that the juveniles scored a larger proportion of 'near-misses' than the adults. It will be convenient for the time being to ignore the outstanding greenfinch group (aviary-reared) and consider the hand-reared birds as the standard group of juvenile greenfinches. Considering only the first four groups in Tables 1 and 2, therefore, the adults and juveniles were roughly equal in their capacity to ignore the black lid.

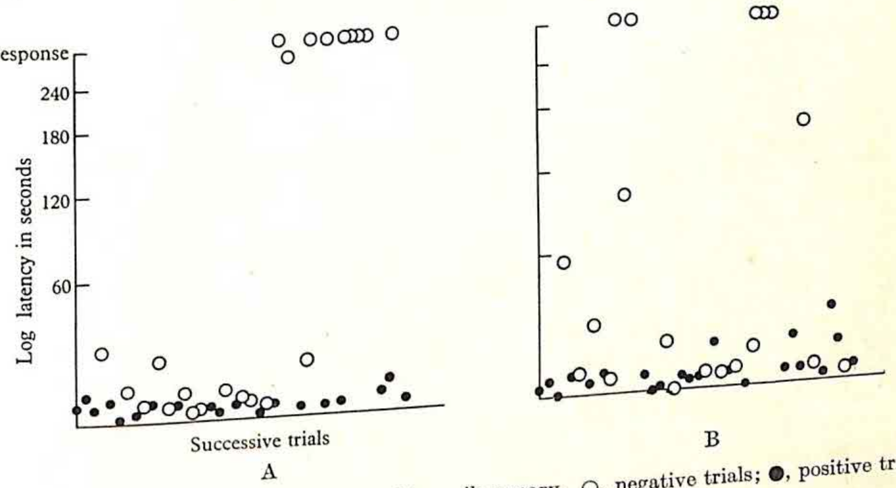
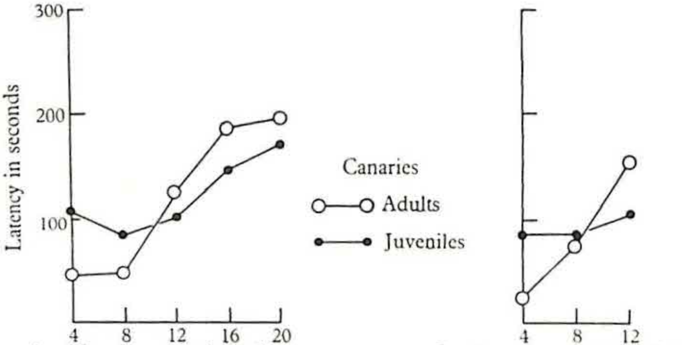


Fig. 2. A, Record of adult canary. B, Record of juvenile canary. ○, negative trials; ●, positive trials.

Table 2

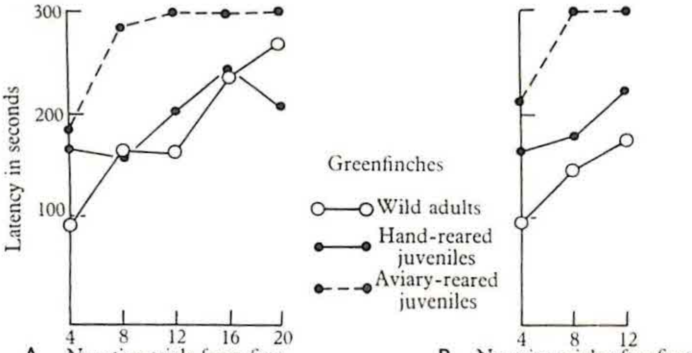
	Canaries		Greenfinches		
	Adult	Juvenile	Adult	Juvenile hand-reared	Juvenile aviary-reared
Mean latency in seconds for twenty negative trials	117 (five birds only)	121 (six birds only)	183	194	266
Percentage correct trials (no response in 5 min.)	41	28	55	51	85
Mean number of negative trials before first correct (no response in 5 min.)	9·1	3·25	5·7	3·3	2·3

In one respect, however, the performance of the juveniles appeared to be superior to that of the adults. It may be seen from the individual records in Fig. 2 that the juvenile canary completed a correct negative trial (its fifth) earlier than the adult shown, whose first correct negative response was achieved only on the thirteenth negative trial. The mean numbers of negative trials before one was completed correctly for the four groups are given in Table 2 and the same difference holds here. To take individual cases, only one out of the eight juvenile canaries failed to score a correct trial within five negative trials whereas only three of the adults succeeded in doing so. The difference was less between adult and hand-reared greenfinches. This difference appears also in Fig. 3A where the average latency in successive negative trials is presented. Here the curves have been smoothed by averaging the scores in groups of four trials. It may be seen that, in the earlier trials, the younger greenfinches and canaries were superior to the adults of their own species; the differences, however, are small.



A. Negative trials from first presentation of black lid

B. Negative trials after first response to black lid



A. Negative trials from first presentation of black lid

B. Negative trials after first response to black lid

Fig. 3. Latency of response to black lid in twenty negative trials.

The results given in Fig. 3A could have been biased owing to another difference between old and young birds mentioned earlier; the latency of response in the first negative trial was longer in juveniles than in adults (Table 1) and, in a few individuals, the black lid was not removed until the second or even the third negative trial. The lid-removing response was, in fact, generalized to the black lid more slowly in the young birds. This effect could bias the results in Fig. 3A in favour of the juveniles and the data are therefore plotted again in Fig. 3B beginning only from the first trial after that in which the black lid was first removed by the bird. The curves in Fig. 3B are similar to those in Fig. 3A; nevertheless, it is not possible to distinguish between an increase in latency due to a failure to generalize rapidly, and an increase due to the lack of reinforcement in negative trials. The possibility remains therefore that, in the early stages, the differentiation may have been easier for the young birds. Elsewhere than in Fig. 3B these initial correct negative trials have been included in the data.

Any superiority of the juveniles was in any case short-lived. Fig 3A shows that the performance of the adults was, on the whole, superior to that of the juveniles in the later trials.

This difference was produced by a lack of stability in the negative response, more marked in the juveniles than in the adults. The juveniles showed early signs of making the differentiation, but it later broke down, reappeared, broke down again, and so on. Fig. 4, where the number of correct negative responses is plotted as a percentage of the whole, shows the extent to which the early improvement of the juveniles fell off in the second half of the test period. It is true that complete stability, consisting of long negative and short positive latencies, was achieved by very few birds during the test period. The total number of incorrect positive responses for each group is given in Table 1 (11), and the average positive latency for each group in Table 1 (12). Consideration of individual performance on both positive and negative trials is important at this point. If five consecutive correct negative trials is taken as a criterion of stability then, in the canary groups, this was achieved by four of the seven adults and one of the eight juveniles. This latter bird, however, always gave very long positive latencies and it alone accounts for

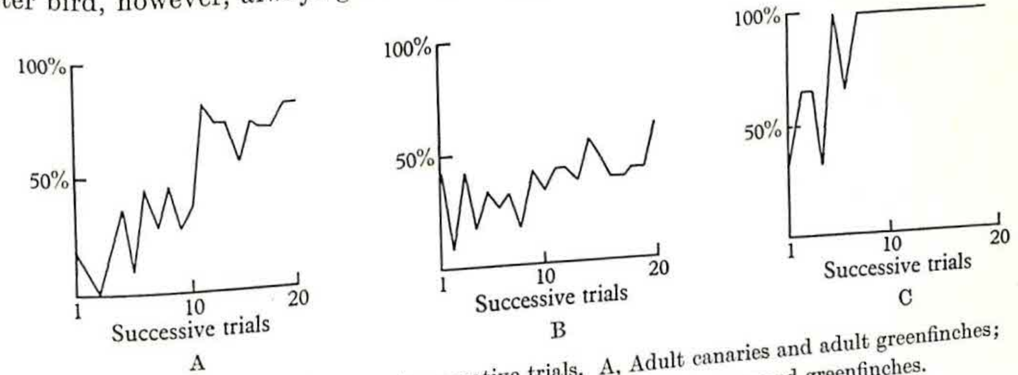


Fig. 4. Correct responses in successive negative trials. A, Adult canaries and adult greenfinches; B, juvenile canaries and hand-reared greenfinches; C, aviary-reared greenfinches.

half the incorrect positive responses for this group. In the adult group, positive responses were stable except for one bird which achieved the criterion of stability at the expense of all the incorrect positive trials for this group (in Table 1 (12) average positive latencies are given with and without the data from this bird). Therefore three of the adult canaries achieved a stable overall performance whereas none of the juveniles did so. Amongst the adult greenfinches, all positive responses were rapid and two birds reached the criterion of stability in the negative response (the other two failed to do so, by one trial only). The two birds in the juvenile group which also achieved this criterion did so only at the expense of considerably lengthened positive latencies and one incorrect positive response each. Again, a stable performance was obtained only from the adults.

Results obtained from the aviary-reared juvenile greenfinches were quite clear-cut. This group was tested at ages comparable with the hand-reared group, but their performance was, in every respect, superior. Correct negative responses were established very rapidly and remained remarkably stable (Figs. 3A, B; 4). Positive latencies were short and stable.

The aviary-reared birds thus gave a performance which was superior to all individuals in all other groups, but the other differences were less clear-cut. It is, however, true to say that none of the other juveniles reached a level equal to that of the best (roughly half) of the adults of their own species. This point is of some importance in considering the number of birds which, for one reason or another, were not tested. The adults were more difficult to train and were, for that reason, a more highly selected group than the juveniles. If the birds which failed to remove the lid did so because they were, in general, 'poor learners', this will have biased the results in favour of the adults, but does not alter the fact that some adults achieved a stable performance whereas no hand-reared greenfinch or young canary did so. What evidence there is, however, suggests that those birds which could not be trained adapted most rapidly and most stably to the various elements in the experimental situation. One canary was abandoned because it would remove a lid placed so that the bait was visible, but not when it completely covered the bait. In other cases, the birds removed the lid on at least one occasion, but with a snatching or fluttering movement so that the lid flapped at a time when it was still a relatively strange object, causing the bird to fly up without feeding. This type of incident seems to have produced a very stable avoidance response in some adults as the untrainable birds frequently gave up feeding even when the bait was placed in a dish with no lid. Some further evidence bearing on this point has been obtained more recently from wild great tits tested in the same situation at different ages. The numbers here are small but training has become more difficult as the birds approach the age at which the differentiation becomes more stable. The results in the present experiment should not be biased by the other cause of wastage, the failure to generalize the lid-removing response to the black lid, as this affected one bird in each group of greenfinches.

The results suggest that the differentiation may have appeared rather more rapidly in the young than in the older birds. However, the adults tended to give a more stable performance. In a group of juvenile greenfinches reared by their parents in an out-door aviary the differentiation appeared very rapidly indeed, and also remained stable.

IV. Discussion

In the earlier experiment (Vince, 1958) it was found that behaviour comparable with the extinction of conditioned reflexes remained far less stable over a number of days in young as compared with adult birds. In the present experiment, the situation required the establishment of a differentiation. This situation was chosen because, according to the Pavlovian theory, the same process of internal inhibition is involved in a differentiation as in extinction. We are concerned with the question of whether changes in behaviour following a lack of reinforcement (internal inhibition) are effected by age and experience.

The results support the hypothesis stated earlier: that internal inhibition is unstable in young as compared with adult birds. In neither group of young canaries nor in the group of hand-reared greenfinches was a stable differentiation established without instability in the positive response, whereas this did occur in about half of the adults tested. As in the previous experiment, however, the results obtained from young greenfinches reared out of doors showed features similar to both adult and juvenile groups. In the present experiment, longer latencies in the first few trials occurred in all juveniles, but the stability of negative responses in the aviary-reared birds was comparable with, in fact far superior to, that of the adults.

It is clear that these results cannot be understood in terms of age alone; as in the previous experiment, the results obtained from aviary-reared greenfinches emphasize the importance of early experience in the type of development under consideration.

The results may be interpreted as follows: the canary data suggest that there may be a process of development in internal inhibition from unstable to more stable as a result of age. We have, however, observed the canaries at only two stages of development. The greenfinch results may well give rise to the speculation that this process of development must be represented by a curve. Of the two groups reared in an out-door environment, it seems reasonable to suggest that the adults will have already passed through the stage of rapid and stable adjustment manifested by the juveniles. If this is so, there seems to be a weakening of this type of negative response between the juvenile and adult stages. Again, given their richer environment, it seems likely that the stage of development represented by the aviary-reared juveniles is in fact more advanced than that of the hand-reared birds, although the ages at which they were tested were comparable. If this is so, it gives a rising curve of development during the juvenile stage.

The number of juvenile aviary-reared greenfinches tested was small, but this group gave the most homogeneous results and were most markedly different from all other groups, thus strongly suggesting that there is a peak in this type of development during the juvenile stage. The remaining groups were more variable, but their performance has been analysed in some detail because it suggests that, although the overall differences were slight, the behaviour of the adults was rather different from that of the juveniles in this situation. This difference would seem to point to a process of development beyond the juvenile stage, rather than regression in the adults towards the behaviour of the very young birds.

The foregoing interpretation is speculative, but it suggests an hypothesis for further experiment: that in an un-reinforced situation, internal inhibition in birds develops during the juvenile stage (strengthening or making for a more precise adaptation) and subsequently changes again, being manifested more slowly, but remaining stable. The added hypothesis that the shape of the curve representing internal inhibition depends on the conditions of rearing—becoming steeper with a more varied environment—is still more speculative, but current experiments on more homogeneous groups of birds reared under wild or caged conditions support this suggestion.

There is a little evidence from other sources which gives some support to these hypotheses. Experiments have shown that internal inhibition as measured in a differentiation is weak in the very young and later strengthens (Panchenkova, 1956, in rats, and unpublished experiments by the writer, in chickens). There is also evidence that variety of early experience is essential for normal adult behaviour in dogs (Thompson & Heron, 1954*a*, *b*): dogs reared under very restricted conditions later tended to perseverate and were incapable of delayed reactions. If this work is considered in the terms used above, it would seem that, in dogs, inhibition fails to develop normally under restricted conditions of rearing.

Very little is, in fact, known about the type of development under consideration and the experiments described are merely exploratory. We do not yet know what are the important variables in this type of investigation. Age has to be considered, but the other variable, type of early experience, might be resolved into a number of factors. For the time being, it is proposed to consider variety of experience, i.e. range of opportunity for activity or adaptation, as the significant factor here, and the work of Thompson & Heron supports this.

This work was carried out with the support of the Medical Research Council.

REFERENCES

PANCHENKOVA, E. F. (1956). The ontogenetic development of conditioned reflexes in the white rat. (In Russian.) *Zhurnal Visshei Nerv'noi Deyatelnosti*, **6**, 312–8.

PAVLOV, I. P. (1927). *Conditioned Reflexes.* London: Oxford University Press.

THOMPSON, W. R. & HERON, W. (1954*a*). The effects of early restriction on activity in dogs. *J. Comp. Physiol. Psychol.* **47**, 77–82.

THOMPSON, W. R. & HERON, W. (1954*b*). The effects of restricting early experience on the problem-solving capacity of dogs. *Canad. J. Psychol.* **8**, 17–31.

VINCE, M. A. (1958). 'String-Pulling' in Birds. II. Differences related to age in greenfinches, chaffinches and canaries. *Anim. Behav.* **6**, 53–59.

(*Manuscript received* 3 *February* 1958)

EFFECTS OF TELEVISION ON THE INTERESTS AND INITIATIVE OF ADULT VIEWERS IN GREATER LONDON

By WILLIAM A. BELSON

*Research Techniques Unit, London School of Economics**

1. *Aims.* The aim of this study was to measure the effect of television upon the interests and the initiative of adult viewers in a specific urban area and to relate the extent of the effect to length of set ownership.

2. *Methods.* The Stable Correlate design was used for the isolation of effects. There was extensive preliminary test construction, and the measurement of television's effects was made against a representative cross-section of the public's *interests* in the one case, and of their *acts of initiative* in the other. The final survey was based upon about 800 cases and the construction of measuring devices upon about 3000.

3. *Results.* The effect of television has been to reduce both interests and initiative. The reduction of interests is not only in terms of activity level, but in viewers' *feeling* of interest as well. Even when interests are featured in television programmes themselves, the loss to television is made up to only a very small degree. Television's effects do, however, vary markedly from one group of interests to another. The loss in initiative and in interests extends over a period of five to six years. Generally speaking, the loss is greatest in the first few years, after which there is a gradual recovery.

I. INTRODUCTION

When British television broadcasting started again in 1947, less than 1% of the adults in the United Kingdom were viewers. By the end of 1955 the figure was nearly 40% (Emmett, 1956) and this had risen to about 60% in early 1958. At that time, more than 95% of the population lived in areas where television programmes could be received. Moreover, more than half of those with sets could receive commercial television as well, and that proportion was steadily growing. Table 1 shows some of the details. Nor was this a coverage in name only, for the average viewer watched evening television for nearly 40% of the available viewing time, and hundreds of thousands watched it for nearly 60% or more of that time.

Table 1. *The growth of television*

Year ending	Number of transmitters	Percentage in areas able to receive	Percentage with TV sets	Percentage with sets able to get ITV
1950 (Dec.)	2	42·3	4·3	—
1951	3	65·1	8·7	—
1952	5	80·9	14·0	—
1953	9	86·3	21·8	—
1954	12	90·7	30·6	—
1955	15	93·7	39·8	4·5
1956	19	95·2	49·4	19·7
1957	23	97·0	58·4	32·5
1958	29	98·3	65·0	45·0

In the circumstances, it is inevitable that speculation should have grown about television's impact on society. It was said to have done many different things. On the negative side, it was said to have reduced attendance at the cinema and at sporting events; to have caused people to read less; to be 'killing' the population's interests; to be reducing initiative, imagination, and creative activity. On the other hand, there were claims to the effect that it was bringing the family together again; that it was reducing irritability and taking some of the boredom out of day-to-day existence. It was said to be educational and informative and to be broadening the cultural horizons of the nation. There were many more such theories, but between them they made of television both a menace and a thing of promise.

---

* The writer is Research Fellow in Survey Psychology in the London School of Economics. Previously he was Senior Psychologist in the Audience Research Department of the B.B.C. where he worked in communication research and on methodological developments. The inquiry reported here had financial assistance from the Research Fund of the University of London and from the B.B.C.

In America, speculation of this kind has been followed fairly rapidly by research. There have been studies there of television's impact on: family life (Coffin, 1948; Alldredge, 1950; Fine, 1952; Lawton, 1955); on cinema-going, sport (Jordan, 1950; Sheatsley and Borsky, 1953) and leisure-time activity generally (Batten, Barton, Durstine & Osborn, 1951); on reading, sale of products (N.B.C., 1952) and on sociable behaviour (Riley, Cantwell and Ruttiger, 1949; Meyersohn, 1953 and 1957; Hamilton and Lawless, 1956.) Meyersohn's annotated bibliography of 1953 brought together most American studies up to 1955. His 1957 paper summarized research up to 1956.

But here in Britain there have been relatively few checks upon the many and varied theories which exist. This, however, is a report on one of them, carried out by the writer over a period of four years. It is primarily a study of television's effects on viewers' interests considered at both the active and the passive level. Built into it, however, is a separate inquiry, concerned with television's effects on initiative (defined as 'the ability for independent action or original conception' after Warren, 1934). The study was made in Greater London.

II. METHODS

Design

Anyone working in this field is likely to encounter a serious problem when it comes to isolating the effects of television from all the other things which can be mistaken for effect (McPhee, 1953). Some investigators have partly met this difficulty by studying the behaviour of people before and after they get sets, using a control group of non-viewers to tap the influence of purely extraneous events. An important theoretical difficulty here is that, unless the control group is very carefully matched to the viewing group, the influence of extraneous events on the control will *not* necessarily provide a measure of how those extraneous events affected the *viewers*. There are also problems of a practical kind, one of which is that the method is extremely wasteful: this is because so many people have to be tested *before* the coming of television to ensure a reasonable sample of actual buyers of sets. Then there is the problem of wastage of even these, and of the bias to which this can lead.

Other inquiries have been little more than comparisons of viewers and non-viewers, differences between them (e.g. in attendance at music halls or at sporting events, in time of retiring at night) being interpreted as effects of television. In such cases, the obvious objection is that viewers and non-viewers may have been different to start with, and that present differences between them can give a seriously misleading picture of television's effects. Nor is it usually sufficient simply to match viewers and non-viewers in terms of age, sex and social class, in the expectation that this will rule out such original differences or substantially control them. Matching can be effective only if the criteria in terms of which it is done are *empirically* selected as the most relevant ones to use.

The Stable Correlate method. The research design which was used in this inquiry is in fact a development of matching, the matching variables being selected empirically, and there being no discarding of unmatched persons. For reasons which will become clear in the following paragraphs, the writer has called it the Stable Correlate method (Belson, 1956*b*).

The method *begins* with the simple comparison of viewers and non-viewers in terms of whatever is being studied as possibly affected by television (e.g. amount of cinema-going, frequency of occurrence of sociable acts or of acts of initiative, range of ambitions). After this, the non-viewer assessments are adjusted to provide as accurate as possible an answer to the following question: '*What would the average non-viewer assessment have been had these non-viewers been the same* (*to start with*) *as the viewers in terms of the characteristics which matter, i.e. in terms of the characteristics which are predictive of or correlated with the behaviour or ability being studied?*' Once these correlates or matching variables have been selected or developed, the adjustment can be achieved through a regression equation or, better, through simple weighting. After this adjustment to the non-viewer assessment has been made, the remaining difference between this adjusted score and the *viewer* score can be interpreted as giving a much closer estimate of television's effect than is obtainable through unadjusted comparisons or through *a priori* matching. And it is quite usual for the estimate to prove very close indeed.

The success of the Stable Correlate method turns upon the development of effective correlates or matching variables. The usual approach is to include in the testing programme (or in the questionnaire) some 50–100 possible correlates of the variable under study. In the check on television's effects on interests, some of the 146 proposed predictors were: occupational background; type of family; the presence of potential child-minders in the household, or nearby; whether or not the respondent had ever passed an examination or gained a certificate of some kind; the nature of the public transport available; the ownership of various durable goods; past activities and achievements of a special kind. The measurement of the variable under study having been made (i.e. the criterion score having been established) there follows a form of mass item-analysis to establish the predictive power of the many proposed correlates. Using the Wherry-Doolittle formulae (Garrett, 1947), this can lead to the selection of that three or four of them which, taken together, give as good a prediction of the criterion score as all of them taken together.

This selection of predictors must, however, be made through the *non-viewer* scores, for these are the scores which are to be adjusted through the predictive system so developed.

The third step in the Stable Correlate method is to break the non-viewer record cards into as many subgroups as the 3–4 matching variables provide for. Thus, three matching variables, each giving a two-way split, will provide for eight distinct subgroups. Four of them can lead to sixteen subgroups. The average criterion score for each of these is calculated and this is multiplied by the number of *viewers* in the equivalent subgroup. These subproducts are added and divided by the total number of viewers. The result represents the *adjusted non-viewer score*. In other words, it represents what the non-viewers would have scored had they been the same as the viewers to start with. The whole adjustment process could of course have been made through a regression equation, but this is tedious and less accurate than the 'multiplication' or weighting method described. The adjusted score can then be compared with the unadjusted viewer score, and the difference interpreted as an estimate of the effect of television.

One point is vital, however. The matching variables used must not themselves be open to change by television, for otherwise the adjustment to the score of the non-viewers could only lead to a whittling away of the evidence of real effect. In other words, the matching variables must be *stable* in this sense (i.e. *stable correlates*).

Two final points about the method should be made. One is that in practice it is rarely possible to develop a multiple predictive power (in the matching variables) of more than ±0·8. Naturally this leaves open the possibility that, after matching, the viewers and non-viewers still differ in terms of some relevant characteristics. Accordingly, it is necessary to build into the design a device which allows tests to be made of the degree to which the matching has in fact been complete. This device in fact involves the matching of selected subgroups of the *non-viewers* in the expectation that a completely adequate matching procedure will remove *all* differences between them in terms of the criterion score. This aspect of the method has been described in detail elsewhere (Belson, 1956*a*).

The second point to be made is that for each variable being studied for change, a fresh search for matching variables must be undertaken. Thus, in the present study, the matching variables for measuring television's effects on initiative are substantially different from those used in the study of interests.

Making the necessary measuring devices

The various phases of test-construction were extended over a period of two years and involved, in all, about 3000 people. Details are available in the full report of the inquiry (Belson, 1957), but the following short description gives the essential features of the construction process.

Interests were studied through a 50-item cross-section of the interests of those in the survey universe (Greater London). The development of this sample was empirical, and it was based on survey methods.*

---

* This was done in three stages. In the first stage, a sample of 300 people took a week in which to list their interests, and these were used to develop an inventory of the interests existing in the survey area. The 500 different *kinds* of interests so developed were then made the subject of a series of surveys to establish their relative followings in the survey population. From this weighted list of interests, each with the size of its following against it, a sample was then drawn on the principles ordinarily used to select area units for stratified sampling.

It was deliberately a cross-section, in order to allow the making of a reliable generalization about interests as a whole. There follow some examples of the interests in this sample: 'interior decoration of the home', 'reading books', 'going to the cinema', 'variety and vaudeville shows', 'visiting people', 'strolling in the park', 'hearing military bands', 'international and world affairs', 'going to see sporting events', 'fashions to-day', 'modern developments in science', 'looking at paintings', 'ice or roller skating', 'bird watching', 'collecting things'. The full list of fifty is incorporated in the Appendix.

In a sense, this cross-section constituted a frame-work against which to measure change in viewers' interests. Two measures of interest were used. One was a simple rating on a six-point scale* of the degree of interest which the respondent felt towards some stated activity or thing (i.e. strength of interest), the fifty ratings being aggregated for a total score. The other was a measure of the respondent's *activity* in relation to the fifty fields of interest in the sample. Thus, for each of them, the respondent was asked if he had done some specific thing in the last two weeks and, if so, how often. (E.g. 'Did you go to the cinema in the last two weeks?...If yes, how often?...) There were fifty of these questions and their aggregate was referred to as 'activity level'.

The check on initiative was, in principle, the same. The measure was made in terms of frequency of occurrence of 'acts of initiative', 'initiative' itself being defined formally as the 'ability or capacity for original conception and independent action' (Warren, 1934), or, in popular language, as something 'done off your own bat'. After considerable preliminary work, a cross-section was developed of the acts of initiative ordinarily occurring in the survey universe. As in the selection of the sample of *interests*, this development involved two preliminary surveys. It was followed by item analysis and a system of weighted selection. There were thirty items in the final check-list and they were of the following kind.

When something broke, or ceased to work, I mended it.

When the proper tool or instrument or material was not available, I made do with something else. (For example: I made a rough-and-ready copy of a missing tool or instrument; I thought of a way to make another tool or instrument to do the same job; I thought of a way to use different kinds of material to do the same job; and so on.)

I started up a conversation with someone I didn't know.

When my transport was running late I used another means of transport.

With each of the thirty, respondent was asked if it occurred during the 'past two weeks' and, if so, to select a statement indicating the degree to which it was his/her 'own idea to do it'. An aggregate score was developed.

The survey in which these measuring devices were used was spread over a period of about a year (to avoid seasonal bias) and was administered through a questionnaire of a new kind, designed to meet the special needs of such administration. The survey area was Greater London and the sample itself consisted of a 450-person cross-section of *viewers*, plus a 350-person cross-section of non-viewers.

III. FINDINGS

The stable correlates selected

The matching criteria selected for the study of initiative and of the two aspects of interest are set out in Table 2. The 'past acts' named in this table are various composites of past ownerships and of past achievements relating to the period before 1950. As a composite, each has in it about thirty different items, though the items in each tend to be different. The fact that the second and the third predictors make very little difference to the predictive power of the first, is a function of the heterogeneity of that first predictor. Ordinarily, the second and third variables make a greater contribution. They were none the less included here to give greater stability to the matching procedure.

---

* This six-point rating scale was selected from a battery of 27 scales. It was, in fact, that one amongst them which gave a consistently high prediction of score based on the whole 27 of them. The 27 scales, taken together, approximated to the universe of dimensions to which the public refers when they judge that someone is 'interested' in a particular thing. The scale selected can, therefore, be regarded as an index of 'interest'.

Table 2. *The stable correlates of interest and of initiative*

Variable under study	Stable correlates selected (in order of priority): Variable	Stable correlates selected (in order of priority): Correlation	Multiple correlation achieved: First 2	Multiple correlation achieved: All 3
Interests, considered generally				
(*a*) Strength of interest (as felt)	Past Acts (C1)*	+0·61	0·615	—
	Age	+0·13		
(*b*) Associated behaviour	Past acts (C2)*	+0·66	+0·685	+0·688
	Examinations passed	+0·20		
	Authority in job	+0·12		
Initiative	Past acts (C3)*	+0·77	+0·783	+0·784
	Has child aged 16 or less	−0·13		
	Age	−0·19		

* C1, C2 and C3 are separate composites of past ownerships/achievements, each including about 30 items and each developed through item analysis.

Table 3. *Effects of television on viewers' interests and initiative**

Characteristics studied	Non-viewer scores: Before matching to viewers	Non-viewer scores: After matching to viewers	Viewer scores	Loss or gain	Significance of difference (*P*)
Interests					
Strength of interest	89·08†	88·14	84·19	−5%	0·01
Level of activity	18·32	17·80	16·07	−10%	0·01
Initiative	7·40‡	7·48	7·05	−6%	0·01

* 800 persons, randomly distributed. † Max. score = 200. ‡ Max. score = 30.

Television's effects on viewers' interests

In studying the effects of television upon interests, viewing as such, considered as an interest, is being excluded, for this is a study of the effect of viewing television upon the viewer's other interests. The main results are set out in Table 3, which also includes some details of the matching process. Thus, it shows that the different adjustments led to various shifts in non-viewer scores, moving them nearer to viewer scores in two cases and further in the other case. From this table it will be seen that there has been a statistically significant reduction in viewers' interests in terms both of strength of interest (a 5% loss) and activity level (a 10% loss); and this applies too with the frequency of occurrence of acts of initiative (where the loss is 6%).

It is important to know, however, how permanent these losses are and at what time after buying a set they are at their greatest. Since 'length of set ownership' was known, the effect of television after different periods of ownership was also derivable. This required that the matching procedure be applied separately for each of the 'set-age' groups. Thus non-viewers were first matched to the viewers who had owned sets for *one year or less*, and an estimate of television's effects after that length of ownership was derived. The non-viewers were then matched *afresh* to the viewers who had owned sets for *one to two years*, to derive an estimate of effect, after *that* length of ownership. Similarly, with the 2+ to 4 year group, and for those with sets for 4+ to 6 years. By linking these points, it is possible, *granted a reasonable matching* of these viewer groups, to derive trend curves of the kind shown in Fig. 1. What this suggests is that activity associated with the viewers' interests undergoes a sharp reduction in the first year of set ownership (14·0%) and that this degree of loss is maintained during the second year (13·9%);

thereafter, however, there is a steady tendency towards recovery (9·7 % after 2+–4 years), and after 4+–6 years, activity associated with interests is almost back at its pre-television level.

With strength of interest—that is, a feeling of interest, irrespective of whether or not it is accompanied by activity—the position is somewhat different. In the first year there is no significant change; during the second, however, there is a fall by 8·6 %. After that comes a tendency towards recovery: the loss is 5·9 % after 2+ to 4 years, and 3·3 % after 4+ to 6 years. The finding that there is no fall in strength of interest during the first year—despite a sharp reduction in associated activity—seems to have its explanation in the fact that television programmes have, quite naturally, featured many of the interests in the sample at some time or other in the past. Further reference will be made to this a little later on.

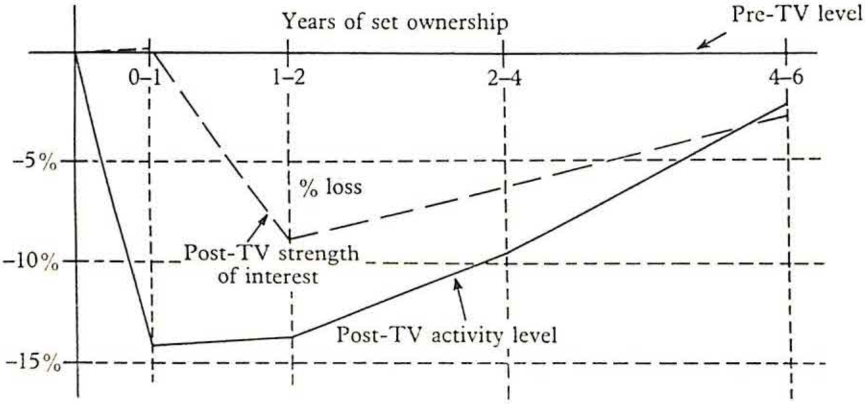


Fig. 1. Television's effect on interests and associated activity according to length of set ownership.

Interestingly, the patterns of change vary with the frequency with which the interest is given active expression. Thus the number of interests which occasion activity fairly frequently is reduced by about 20 % in the first year; this is followed by a partial recovery in the second year, but even at the end of about 5 years, the loss is still at the 10 % level, and there is no sign of its becoming less. On the other hand, the number of interests occasioning activity at a lower rate is reduced less sharply in the first year, continues to decrease in the second year (12·1 %), but thereafter undergoes a full recovery. The pattern with 'strength of interest' is much the same. Details about changes of this kind are given in Fig. 2.

What all this means is that, from a long-term point of view, the main loss is in viewers' more intensely pursued interests, the recovery with those pursued *less* intensely being complete. Another way of putting this is to say that there is in the long-term a full recovery in the diversity or variety of behaviour associated with viewers' interests, but at a somewhat lower level of frequency. The position is much the same with the intensity of viewers' *feelings* of interest in various things.

Looked at from the long-term point of view, one might also say that television makes a difference at first but that, given time, people can 'take it in its stride'—as, indeed, they appear to have done with many an earlier innovation. From the short-term point of view, however, the position is that there is an appreciable reduction in viewers' interests during the early years of set-ownership, and that though recovery tends thereafter to occur, the pre-television level is closely approached only after about 5 years. And if the buying of television sets continues at its present rate, it is likely that for society as a whole there will be an appreciable reduction in interests for some years to come.

Quite apart from the question of the size and the pattern of the loss, it is clear on the evidence that one of television's effects has been to render interests more passive, *for the loss in behaviour is greater than that in 'strength of interest'*. This passivity is specially marked in the first year, where the loss is *only* in terms of behaviour.

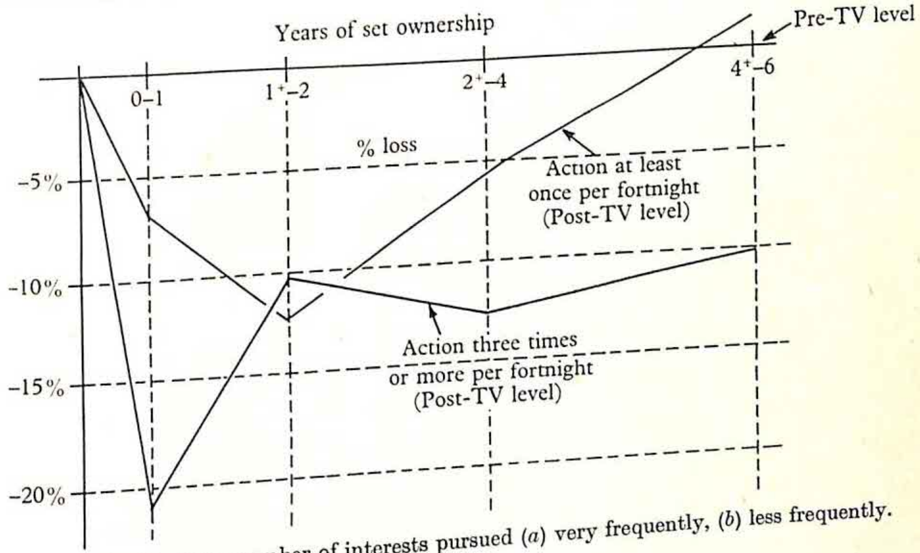


Fig. 2. Television's effects on number of interests pursued (*a*) very frequently, (*b*) less frequently.

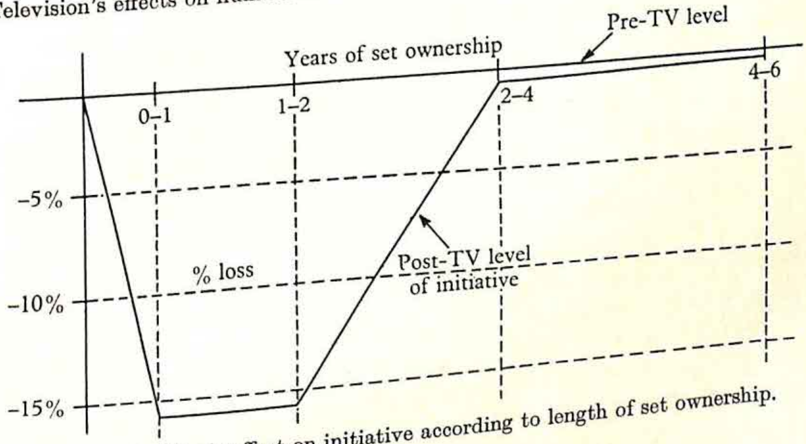


Fig. 3. Television's effect on initiative according to length of set ownership.

Television's effect on viewers' initiative

The same sort of analysis was made with initiative—with substantially the same result. (See Tables 2 and 3 above.) There is a sharp reduction in the first year (15·8%), which is maintained in the second (15·7%). During the third and fourth years, however, there is a relatively complete recovery which is thereafter maintained.

Changes when interests are featured in television programmes

It is an essential feature of the television scene that some interests are featured in programmes, and that others (like visiting, being at home, playing darts, crossword puzzles, membership of clubs and associations) tend not to be. It is understandable that there should be some reduction with the latter group. But what happens with those that are directly featured? To what extent does this produce or enliven viewer interest in those things?

In addition to the check through the 50-item cross-section of the population's interests, a further study was made of a cross-section of those interests which had been featured in television programmes.

Television's effect on these is only slightly different from the pattern of effects already presented: an early loss, followed by recovery.

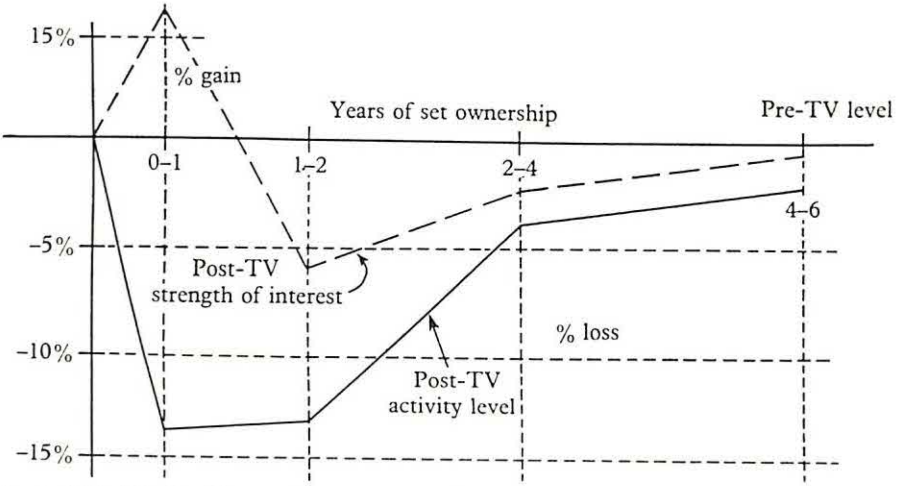


Fig. 4. Television's effect on interests when these are featured in television programmes.

There are two things to be noted about these trends, however. One is that, despite a later decrease, 'strength of interest' *increases* during the first year. Presumably this increase is a novelty effect, though it is essentially a passive increase for it is accompanied by the usual sharp loss in associated activity. The second point to notice is that the loss tends to be slightly less than with the general sample and, by implication, less than the reduction with interests which are *not* so featured. The fact that the overall picture is one of reduction should not, it is suggested, be interpreted to mean that merely to treat an interest in television programmes is to reduce it. There is, in fact, evidence that with some interests this happens, but a more likely interpretation is that treatment of interests in television programmes makes up to only a small degree the reducing effect which the process of continued viewing exerts on such interests in that longer period when they are *not* being so featured. It is as though the television medium carries a basic handicap, such that the person who wants to use it to develop interests of one kind or another has to start from 'well behind scratch'.

Television's effects on a number of specific interests

Because this inquiry was based upon samples of the population's interests, it was possible to arrive at conclusions about television's effects on interests as a whole. However, these conclusions inevitably obscure much about the way in which the many specific interests making up the samples were affected, and tell relatively little about the dynamics of the medium.

Information about these could, however, be derived by applying the stable correlate method to each of them separately in order to establish television's effects. This was done for eighteen of them and the results are shown in Table 4. Several of these results deserve comment, in some cases because of their importance as social issues, and in others because they seem to bring some of the special features of television's impact.

Table 4. *Television's effects in various fields of interest: showing change in 'strength of interest' and in 'associated activity'*

Interest studied	Effect of TV % Change	Effect of TV P
Cinema-going		
Interest in going	−20·6	0·01
Attendance	−33·1	0·01
Theatre-going		
terest in going	−9·1	0·01
Attendance (at play)	−43·7	0·01
Ballet		
Interest	−11·2	0·04
Attendance	−21·2	0·01
Looking at paintings		
Interest	−7·0	0·09
Attendance at gallery	+52·1	0·01
Politics in Great Britain		
Interest	−6·1	0·11
Reading about it	−12·8	0·01
Membership of associations/clubs		
Interest	+0·8	NS*
Attendance	+1·7	NS
Book reading		
Interest	−9·4	0·01
Books read	−22·6	0·01
Gardening (indoors or outdoors)		
Interest	−3·9	0·16
Amount done	0·0	NS
Card playing		
Interest	−1·3	NS
Time playing	−10·3	0·14
Playing gramophone records		
Interest	−17·5	0·01
Playing done	−7·0	NS

Interest studied	Effect of TV % Change	Effect of TV P
Events in different parts of world		
Interest	−7·8	0·02
Reading about them	−6·8	0·12
Modern developments in science		
Interest	+4·0	NS
Reading about them	−16·9	0·03
International and world affairs		
Interest	+1·5	NS
Reading about them	−11·9	0·01
Places of historic importance		
Interest	−5·2	0·14
Made a visit	−29·1	0·01
Information about people in other countries		
Interest about them	−5·2	0·09
Reading about them	−26·1	0·01
Going to see sporting events		
Interest in going	−9·2	0·01
Attendance	+10·5	0·14
Major soccer matches		
Interest	+15·6	0·01
Attendance	+18·2	0·01
Horse racing or horse jumping		
Interest	+11·2	0·02
Attendance	+47·4	0·01
Home orientated activities (various)		
Interest	−5·8	0·01
Activity	−3·8	NS
Sociable activities (various)		
Interest	−2·8	0·11
Activity	−7·0	0·01

* Where NS is shown, $P = 0·17$ or more.

Cinema-going. Cinema-going was reduced by approximately 33% (for all viewers), while *interest* in going fell away by about 21%. For most of the interests studied, year by year trends were also available, and for cinema-going these indicated that attendance underwent a moderate reduction in the first year (11%), but a very substantial one thereafter: 42% in the second year, 43% in the third and fourth years, and 30% in the fifth and sixth years. *Interest in going* was down by 16% in the first year, 19% in the second, 26% in the third and fourth years, and 25% in the fifth and sixth years. Compared with other changes, the 'cinema-going' interest is one of the hardest hit by television and one of the least resilient.

Gardening. When the changes for all viewers (irrespective of age of set) are combined, the overall change in the amount of gardening done is nil. This, however, masks what may be a telling trend. Thus a first-year reduction, by about 23%, is followed by a second-year gain of 13%. After that there is a tendency for gardening activity to return to its pre-television level.

Sporting Events. Television led to a sharp increase both in attendance at and interest in major soccer matches. There were similar increases in respect of horse racing and/or show jumping. Though it is not shown in Table 4, there is also evidence that, for viewers, television has increased both activity and the feeling of interest in connection with Wimbledon tennis and skating (conclusions based jointly on the extent of present difference between viewers and non-viewers *and* the nature of the original differences between them). The one oddity about this set of information is that viewers' *feeling* of interest in 'going to watch sporting events' (i.e. sporting events considered generally) was reduced by about 9 %. This may be a function of the unstructured nature of the reference, but it may be something more basic.

Reading. The reading of books for relaxation or general interest was reduced by 23 %, and even the interest which the viewer *felt* in such reading went down by a moderate amount. Along with this, there was a reduction in the amount of reading done in connexion with such things as: 'events in different parts of the world' (7 %); 'modern developments in science' (17 %); 'international and world affairs' (12 %); 'politics in Great Britain' (13 %); 'people of other countries' (26 %). Most of these specific fields of interest have in fact been the frequent subject of television broadcasts, and the point which should perhaps be noted here is that television presentation seems in these cases to be acting as a direct substitute for the *printed* word. Anyone attempting to evaluate the seriousness of this would of course have to keep it in mind that there may be millions of people who did not in any case acquire such information through reading (and perhaps not at all), but who now get it through television. These specific reductions appear generally to be related to newspapers and specialist weeklies or journals. As far as the reading of *books* is concerned, the evidence does not go far in indicating which kinds of books are the more affected.

Membership of clubs and associations. Contrary to speculation, television does not appear to have produced a permanent reduction in viewers' attendance at clubs or associations. There are, however, some sharp variations, year by year, in the interest *felt* in attending: no change in the first year; a large reduction in the second year, followed by a return to the pre-television level and, after about five years, an appreciable gain. The yearly trends for actual *attendance* are not available from this inquiry.

Going to see paintings. During the year of the study, television increased viewers' attendance at exhibitions of paintings. The increase was considerable, and it seems more than likely that it was associated with some of the televized presentations by Sir Gerald Kelly. Certainly the attendances at public galleries after certain of his broadcasts were greatly swollen. The specially interesting feature of this evidence, however, is that there was also an *overall reduction* in viewers' feeling of interest in 'looking at paintings'. This paradox has its explanation in the fact that the increase in attendance was largely on the part of the minority who had been before. Their interest in looking at paintings had also increased. The *reduction* in interest involved the great majority who had not ever been to see paintings (at an art gallery).

'*Home interests.*' Eleven interests which had to do with the home or which suggested attachment to the home were combined for special study. They included 'interior decoration of the home', 'doing odd jobs', 'festive occasions', 'labour-saving gadgets', 'different ways of preparing food'. Taken altogether, there emerged evidence of a small loss in both associated activity and in feeling of interest (Table 4). The reduction in activity in the first year was approximately 11 %, but after that there was a steady return to the pre-television level. Viewers' *feelings* of interest in these things fell by about 6 % during the first two years but subsequently stayed at that level.

At the same time, there is evidence that television's effects on the different interests within this combination vary considerably, so that these overall figures mask a lot of differences. Incidentally, activities associated with 'home interests' are not to be identified with the general run of household activities, many of which certainly do not come into the category of interests.

'*Social Interests.*' A small sample of the more sociable of the interests was considered in the same way. There were ten of them including 'visiting people', 'making new friends', 'membership of clubs or associations', 'playing cards', 'welfare work'. Associated activity underwent a reduction (15 %) in the first year, this being followed by a long-term return to the pre-television level. The average reduction (irrespective of age of set) was 7 %. Feeling of interest in these things appears to have increased very slightly in the first year, fallen during the second, third and fourth years (minus 6 %), and returned to the pre-television level after about 5 years. Here, too, there was a lot of variation in what happened to specific interests within the group. ('Sociable interests' do not necessarily mean 'sociability', of course.)

The lack of generality in television's effects on interests

One thing which these results make clear is that, despite the nature of television's effect on interests *as a whole*, there is considerable variation in what happens to *specific* interests. Some are reduced quite sharply, others remain little changed, while a few undergo increase. Moreover, a reduction in activity level does not necessarily go with a reduction in viewers' *feelings* of interest. Nor is the opposite the case. While it is only partly indicated here, television's effect on these interests, *year by year*, is also quite varied, there being a lot of deviation from the general pattern of early loss and later recovery. From both these lines of evidence, it follows that a knowledge merely of television's effects on interests as a whole does *not* allow us to say what happens in any specific field of interest. Nor does it help much to know whether or not the interest was television-featured, for in both the 'featured' and the 'not-featured' groups of interests, the variability in effect is very marked indeed.

It was only for the eighteen interests in Table 4 that the full matching procedure was carried out. For the remaining forty-two interests included in the various samples, only the unadjusted differences between viewers and non-viewers are presented. These are given in Appendix 1. While, taken along with other information, these differences provide grounds for speculation about the dynamics of television's effects, such speculation is not attempted in this paper because the aim here is to stay as much as possible with a factual presentation. None the less, the application of the stable correlate matching procedure to the remaining forty-two interests and the development from these of a system of hypotheses seems to be the logical follow-up to this inquiry. It should serve to show which general kinds of interest are *more* affected and which *less* affected by television, and it may serve to indicate *why*. In addition it may well indicate what can reasonably be expected when different kinds of interest are presented through this particular mass medium.

Original differences between viewers and non-viewers

Finally, the study provided information about some of the original differences between viewers and non-viewers. The point is basic, for it bears very closely indeed upon the nature of the matching required and upon the interpretation of results. The two samples, it must be remembered, excluded people aged over 65 years and also that 15% of the population living in the most economically depressed of the polling districts.

Age, occupation and education. For the sample as described, there were no appreciable differences in the mean ages of viewers and non-viewers, though viewers tended to be slightly over-represented in the middle age groups. In terms of occupational background (i.e. present job or last job held), there was also no appreciable difference in the mean level, though there was a slight tendency for the self-employed to be over-represented amongst the viewers. For educational level, there were no appreciable differences in age of leaving school, but (for the pre-television period) considerably fewer viewers than non-viewers claimed to have had further education or training since then (28%:40%). There were no differences in intelligence.

Family and mobility. The families of viewers tend more than non-viewers to include at least one person under the age of 15 years (48%:36%). Their houses are slightly bigger, contain more people, and are slightly more crowded. The viewers also tend somewhat more to have relatives living nearby, but even so they are slightly less mobile than non-viewers because of family commitments.

Pre-television activities and ownerships. There was evidence of differences between viewers and non-viewers in terms of pre-television ownerships and achievements. Thus in the period before television, *fewer* viewers than non-viewers: owned sporting gear; were members of a debating society; did some bird watching; attended a performance of ballet; went somewhere to hear light music played; took a holiday in the country; were members of a lending library; owned a tennis racquet; went to see the architecture of church or cathedral; regularly celebrated a birthday; played a sport fairly regularly; defied an authority; helped pay for their own education; went against the advice or pressure of someone in choosing a career or a way of life; took a magazine or a journal dealing with international affairs; played a musical instrument. On the other hand, somewhat *more* viewers than non-viewers: owned a camera or a gramophone; held a driving licence; went to an Ideal Homes Exhibition; had an outdoor garden; bought or owned a labour-saving machine; employed others; made some furniture; bought a home; took the *Daily Mirror* or *News of the World*.

One of the more important of the implications of these findings (about differences) is that some of the variables in terms of which quite a lot of matching has been done in the past tend now to be less serious sources of bias than formerly. These include age, sex, occupational level, age of leaving school. At the same time, differences, sometimes of an appreciable kind are still to be found in other directions: on the evidence given here viewers appeared to be more materialist than non-viewers, expressing themselves somewhat more in terms of material things; the non-viewers appear to have put more weight upon further education and training, slightly more weight upon the esoteric and the artistic, and to be somewhat the more self-sufficient. Differences in terms of family composition and physical mobility also exist, and there is no telling *what other original differences* might show up through further probing at the psychological level.

IV. Seeing the results in perspective

It is well to conclude by stressing some of the limits of this inquiry. The first is that it is based upon the adult population of Greater London, *less* the 'over 65 years' and *less* that 15% of the population living in the more economically depressed of the polling districts. As it happens, disproportionately few of the viewers were, at the time of this study, to be found in these excluded groups. But the main point is that the results for Greater London do not necessarily apply in other urban concentrations or in rural areas or small towns. The second defining point is that the study was one of television's effects on interests *other than the interest of viewing itself*: thus there is a sense in which it *has* been a study of the effects of the *viewing interest* on other interests.

The final point which must be made is that while viewers are willing enough to talk about the *social* effects of television, they tend to think of television first and foremost as *giving them a lot of pleasure*. Were it necessary, there is massive evidence in audience research findings that this is no mere lip-service to the medium. A measurement of specific social changes produced by television can perhaps be made without reference to other of its effects. But any overall *evaluation* of these social changes must be made against a background of the pleasure which television provides.

References and reading list

Alldredge, C. (1950). *Television, Its Effects on Family Life in Washington, D.C.* Washington, D.C.: Charles Alldredge, Public Relations.

Batten, Barton, Durstine & Osborn, Inc. (1951). *What's Happening to Leisure Time in Television Homes?* New York: Batten, Barton, Durstine, Osborn, Inc.

Belson, W. A. (1956*a*). Learning and attitude changes resulting from viewing a television series, 'Bon Voyage'. *Brit. J. Educ. Psychol.* **26**, 31–8.

Belson, W. A. (1956*b*). A technique for studying the effects of a television broadcast. *J. Appl. Stat.* **5**, 195–202.

Belson, W. A. (1957). The effects of television upon the interests and the initiative of adult viewers in Greater London. Ph.D. thesis, University of London.

Coffin, T. E. (1948). *Television's Effect on the Family's Activities.* New York: Hofstra College.

Emmett, B. P. (1956). The television audience in the United Kingdom. *J. R. Statist. Soc.*, A (General), **119**, pt. iii, 284–311.

Fine, B. J. (1952). *Television and Family Life: a Study of Two New England Communities.* Boston University.

Garrett, H. E. (1947). *Statistics in Psychology and Education.* New York: Longman's Green.

Hamilton, R. V. & Lawless, R. H. (1956). Television within the social matrix. *Pub. Opin. Quart.* **20**, no. 2.

Jordan, J. H. (1950). *The Long-Term Effect of Television and Other Factors on Sporting Attendances.* Washington, D.C. Radio-Television Manufacturers Association.

Lawton, S. P. (1955). *When Television Moves In.* University of Oklahoma.

McPhee, W. N. (1953). *New Strategies for Research on the Mass Media.* Bureau of Applied Social Research, Columbia University.

Meyersohn, R. B. (1953). *Television Research: An Annotated Bibliography.* Bureau of Applied Social Research, Columbia University.

Meyersohn, R. B. (1957). Social research in television. In *Mass Culture, The Popular Arts in America.* Ed. by B. Rosenberg and D. White. Glencoe, Illinois: Free Press.

National Broadcasting Company (1952). *Radio's Effective Sales Power.* New York.

Riley, J. W., Cantwell, F. V. & Ruttiger, K. F. (1949). Some observations on the social effects of television. *Publ. Opin. Quart.* **13**, 223–34.

Sheatsley, P. B., Borsky, P. N. (1953). *The Effect of Television on College Football Attendance.* National Opinion Center, University of Chicago.

Swanson, C. E. (1951). Television ownership and its correlates. *J. Appl. Psychol.* **35**, 352–7.

Warren, H. C. (1934). *Dictionary of Psychology.* New York: Houghton Mifflin.

Original differences between viewers and non-viewers

Finally, the study provided information about some of the original differences between viewers and non-viewers. The point is basic, for it bears very closely indeed upon the nature of the matching required and upon the interpretation of results. The two samples, it must be remembered, excluded people aged over 65 years and also that 15% of the population living in the most economically depressed of the polling districts.

Age, occupation and education. For the sample as described, there were no appreciable differences in the mean ages of viewers and non-viewers, though viewers tended to be slightly over-represented in the middle age groups. In terms of occupational background (i.e. present job or last job held), there was also no appreciable difference in the mean level, though there was a slight tendency for the self-employed to be over-represented amongst the viewers. For educational level, there were no appreciable differences in age of leaving school, but (for the pre-television period) considerably fewer viewers than non-viewers claimed to have had further education or training since then (28%:40%). There were no differences in intelligence.

Family and mobility. The families of viewers tend more than non-viewers to include at least one person under the age of 15 years (48%:36%). Their houses are slightly bigger, contain more people, and are slightly more crowded. The viewers also tend somewhat more to have relatives living nearby, but even so they are slightly less mobile than non-viewers because of family commitments.

Pre-television activities and ownerships. There was evidence of differences between viewers and non-viewers in terms of pre-television ownerships and achievements. Thus in the period before television, *fewer* viewers than non-viewers: owned sporting gear; were members of a debating society; did some bird watching; attended a performance of ballet; went somewhere to hear light music played; took a holiday in the country; were members of a lending library; owned a tennis racquet; went to see the architecture of church or cathedral; regularly celebrated a birthday; played a sport fairly regularly; defied an authority; helped pay for their own education; went against the advice or pressure of someone in choosing a career or a way of life; took a magazine or a journal dealing with international affairs; played a musical instrument. On the other hand, somewhat *more* viewers than non-viewers: owned a camera or a gramophone; held a driving licence; went to an Ideal Homes Exhibition; had an outdoor garden; bought or owned a labour-saving machine; employed others; made some furniture; bought a home; took the *Daily Mirror* or *News of the World.*

One of the more important of the implications of these findings (about differences) is that some of the variables in terms of which quite a lot of matching has been done in the past tend now to be less serious sources of bias than formerly. These include age, sex, occupational level, age of leaving school. At the same time, differences, sometimes of an appreciable kind are still to be found in other directions: on the evidence given here viewers appeared to be more materialist than non-viewers, expressing themselves somewhat more in terms of material things; the non-viewers appear to have put more weight upon further education and training, slightly more weight upon the esoteric and the artistic, and to be somewhat the more self-sufficient. Differences in terms of family composition and physical mobility also exist, and there is no telling *what other original differences* might show up through further probing at the psychological level.

IV. SEEING THE RESULTS IN PERSPECTIVE

It is well to conclude by stressing some of the limits of this inquiry. The first is that it is based upon the adult population of Greater London, *less* the 'over 65 years' and *less* that 15% of the population living in the more economically depressed of the polling districts. As it happens, disproportionately few of the viewers were, at the time of this study, to be found in these excluded groups. But the main point is that the results for Greater London do not necessarily apply in other urban concentrations or in rural areas or small towns. The second defining point is that the study was one of television's effects on interests *other than the interest of viewing itself*: thus there is a sense in which it has been a study of the effects of the *viewing interest* on other interests.

The final point which must be made is that while viewers are willing enough to talk about the *social* effects of television, they tend to think of television first and foremost as *giving them a lot of pleasure*. Were it necessary, there is massive evidence in audience research findings that this is no mere lip-service to the medium. A measurement of specific social changes produced by television can perhaps be made without reference to other of its effects. But any overall *evaluation* of these social changes must be made against a background of the pleasure which television provides.

REFERENCES AND READING LIST

ALLDREDGE, C. (1950). *Television, Its Effects on Family Life in Washington, D.C.* Washington, D.C.: Charles Alldredge, Public Relations.

BATTEN, BARTON, DURSTINE & OSBORN, INC. (1951). *What's Happening to Leisure Time in Television Homes?* New York: Batten, Barton, Durstine, Osborn, Inc.

BELSON, W. A. (1956*a*). Learning and attitude changes resulting from viewing a television series, 'Bon Voyage'. *Brit. J. Educ. Psychol.* **26**, 31–8.

BELSON, W. A. (1956*b*). A technique for studying the effects of a television broadcast. *J. Appl. Stat.* **5**, 195–202.

BELSON, W. A. (1957). The effects of television upon the interests and the initiative of adult viewers in Greater London. Ph.D. thesis, University of London.

COFFIN, T. E. (1948). *Television's Effect on the Family's Activities.* New York: Hofstra College.

EMMETT, B. P. (1956). The television audience in the United Kingdom. *J. R. Statist. Soc.*, A (General), **119**, pt. III, 284–311.

FINE, B. J. (1952). *Television and Family Life: a Study of Two New England Communities.* Boston University.

GARRETT, H. E. (1947). *Statistics in Psychology and Education.* New York: Longman's Green.

HAMILTON, R. V. & LAWLESS, R. H. (1956). Television within the social matrix. *Pub. Opin. Quart.* **20**, no. 2.

JORDAN, J. H. (1950). *The Long-Term Effect of Television and Other Factors on Sporting Attendances.* Washington, D.C. Radio-Television Manufacturers Association.

LAWTON, S. P. (1955). *When Television Moves In.* University of Oklahoma.

MCPHEE, W. N. (1953). *New Strategies for Research on the Mass Media.* Bureau of Applied Social Research, Columbia University.

MEYERSOHN, R. B. (1953). *Television Research: An Annotated Bibliography.* Bureau of Applied Social Research, Columbia University.

MEYERSOHN, R. B. (1957). Social research in television. In *Mass Culture, The Popular Arts in America.* Ed. by B. Rosenberg and D. White. Glencoe, Illinois: Free Press.

NATIONAL BROADCASTING COMPANY (1952). *Radio's Effective Sales Power.* New York.

RILEY, J. W., CANTWELL, F. V. & RUTTIGER, K. F. (1949). Some observations on the social effects of television. *Publ. Opin. Quart.* **13**, 223–34.

SHEATSLEY, P. B., BORSKY, P. N. (1953). *The Effect of Television on College Football Attendance.* National Opinion Center, University of Chicago.

SWANSON, C. E. (1951). Television ownership and its correlates. *J. Appl. Psychol.* **35**, 352–7.

WARREN, H. C. (1934). *Dictionary of Psychology.* New York: Houghton Mifflin.

APPENDIX

Differences between the interests of viewers and of non-viewers; the general sample of fifty interests.

Early in the list are those items where the score of viewers is greater than that of non-viewers, and later on come those where the non-viewer score is the greater. The fifty items making up the general sample of interests are numbered 1–50 (though this gives no indication of the order of their presentation to respondents).

Area of interest	Difference between scores of viewers and non-viewers expressed as a percentage of non-viewer score	
	Strength of interest (%)	Activity level (%)
Major soccer matches*	+16†	+18
Horse racing or horse jumping*	+11	+47
1. Ice or roller skating	+26	+67
2. Motor cars	+10	+66
3. Darts	+10	+14
4. Variety and vaudeville shows	+18	+26
5. International affairs*	+2	−12
6. The Royal Family	+1	+8
7. Interior decoration of the home	+1	+37
8. Tennis (Wimbledon)	+13	+25
9. Weaving	+4	+78
10. Membership of clubs or associations*	+1	+2
11. Rifle or revolver shooting	−12	+8
12. Going to see sporting events*	−9	+10
13. The English countryside	−7	+23
14. Animals and birds of the world	−1	—
Reading about them	—	−21
Visiting the Zoo	—	+183
15. Fashions today	−5	+4
16. Looking at paintings	−7	+52
17. Trade unionism	−9	+51
18. Welfare work	−6	+1
19. Different ways of preparing food	−11	+4
Gardening*	−4	0
Cricket	+2	−3
20. Hypnotism	+5	—
21. Labour-saving gadgets and ideas	+7	−8
22. Modern developments in science*	+4	−18
23. Cinema-going*	−21	−33
24. Strolling in the park	−7	−22
Needlework	−1	—
English towns	−8	−24
25. Bird watching	−17	−32
26. Cards*	−1	−10
27. Photography	−5	−27
28. Physical culture	−14	−10
29. Ballet‡	−11	−26
Architecture of churches and cathedrals	−8	−50
30. Politics in Great Britain*	−6	−13
31. Going to the theatre to see a play‡	−9	−43
32. Military bands	−6	−55
33. Listening to light music	−6	−49
34. Playing gramophone records*	−18	−6
35. Chinaware and pottery	−12	−24
36. Collecting things	−26	−10
37. Book reading*	−9	−27
38. Crossword puzzles	−3	−27
39. Visiting people	−12	−14
40. Festive occasions	−3	−11
Jobs of others	−11	−2
41. Talking about dreams	−17	−42
42. Talking about the past	−8	−7
People of other countries*	−5	−26
Events in different parts of the world*	−8	−7
43. Listening to or attending lectures or talks	−10	−21
44. Places of historic importance*	−5	−29
45. History	−14	−32
46. Making new friends	−11	—
47. Ideas and doings of well-known people	−9	—
48. Sculpture	−39	—
49. Railways	−24	—
50. Gardens	−6	—

* Differences in score have been corrected to eliminate 'original differences' between viewers and non-viewers, and hence the percentage shown approximates to an 'effect of television'.

† Where the percentage is positive, viewer score is *greater* than non-viewer score; where the percentage is negative, the opposite is the case.

‡ Only the 'strength of interest' differences have been corrected to eliminate 'original differences'.

(*Manuscript received* 3 *March* 1958)

DEVELOPMENT OF THE PERCEPTION OF SOCIAL DIFFERENCES IN CHILDREN FROM 6 TO 10

By GUSTAV JAHODA

Psychology Department, Glasgow University

A specially devised pictorial technique to test the ability to perceive social differences, requiring a minimum of verbal instruction, was applied to a sample of 179 children systematically distributed according to age, sex and social status of parents. A performance test of intelligence and a vocabulary scale were also administered.

An expected improvement with age was found with the social perception test, and there were significant sex and status differences. Analysis of covariance revealed that, except in the case of lower middle-class girls, all status differences were in fact a consequence of the association between intelligence level and status, disappearing when the former was held constant.

In discussing the children's responses it is shown that an incipient class concept can exist without the facility for giving it any verbal expression. An examination of the ways in which children build up a conceptual framework regarding social differences indicates that occupational divisions tend to come first with all children; later these are linked with differences in wealth, income and in the case of middle-class children to some extent style of life. While a cognitive grasp of social differentiation is primarily dependent upon intellectual level, attitudes and feelings are largely determined by influences emanating from the child's particular social environment.

I. INTRODUCTION

Recent surveys of the literature on social stratification by MacRae (1953) and Pfautz (1953) indicate a shift away from the earlier broad and sweeping generalizations, penetrating as some of them may have been, towards detailed empirical studies. This applies equally to the subjective elements treated in the past under the heading of 'class consciousness'. A concept of this kind does not easily lend itself to an operational approach, for as Ginsberg (1930, p. 536) acutely noted '...it is extremely difficult to say what exactly one is conscious of when one is class conscious'. Hence the more circumscribed aim of this paper is to gain information about the ways people view themselves and their associates within the framework of the class structure.

The most common approach in this kind of study has been to question people directly. Thus, Martin (1954, pp. 54–9) asked his respondents: 'How many social classes are there in this country? Can you name them? Which of these classes do you belong to? What sort of people belong to the same class as yourself? What sort of people belong to the other classes you have mentioned?' A large majority of the sample experienced no difficulty in giving their answers. Nearly all thought in terms of a three-class system, but there was considerable divergence in the drawing of boundaries, and as regards the criteria employed.

Adolescents were the subjects in several studies (Centers, 1950; Himmelweit, Halsey & Oppenheim, 1952; Jahoda, 1953) which also relied mainly on verbal methods. It became apparent that although boys and girls in their early 'teens were able to discuss various aspects of social stratification, a knowledge of the actual vocabulary of social class could not be taken for granted. As an example, the work of Himmelweit and others (1952) may be cited, who asked 13-to-14-year-old boys: 'Do you think that there are social classes in England?' About half the Grammar School boys and two-thirds of the Secondary Modern School boys failed to understand the question. At the same time the results of the other parts of the inquiry led the authors to the conclusion that at their age the boys had already a thorough grasp of the social class system.

Such a knowledge must clearly be the product of a prolonged learning process, but hardly any attempt seems to have been made to investigate its earlier stages. An outstanding exception is the work of Stendler (1949) in the United States, who applied a variety of techniques with children aged 6 to 15. Only one of these, the picture test, will be briefly described, since part of it closely resembles the procedure followed in the present study. Its purpose was 'to find out what symbols of class the children of different grade levels might recognize' (*op. cit.* p. 29). Four sets of pictures, on whose class ratings adult judges had agreed, were presented to the children; they illustrated homes, clothes, recreations and occupations. The question asked about the people who featured in the pictures was 'Do you think they make lots of money, not much money, or in-between?'; with older children the terms 'high in society, low in society, or in-between' were substituted. Children were also invited to give reasons for their answers. The results may be summarized as follows: children's class ratings of the pictures increasingly approximated those of adult judges with rising age levels; and the higher the child's own social class status, the better its performance. It should be noted, however, that the effect of intelligence was not partialled out.

While the investigation as a whole was admirably designed and executed, the part of the picture test referred to is perhaps open to some objections. It would seem that the mode of questioning was too directive and therefore liable to influence the responses. Furthermore, it was assumed that the criterion or dimension of social class uppermost with young children is that of money income. This may of course be the case; but in the absence of independent evidence, the failure of the youngest children to grade the pictures according to this criterion cannot be taken as showing conclusively that they were not yet aware of *any* facet of social class.

The present work is intended as a contribution towards a broader genetic treatment of the problem, along the lines advocated by Mace (1938, p. 155). Casual observation of pre-school children convinced the writer that even at such an early age a rudimentary capacity for social discrimination is already beginning to appear in some children. At this level the apprehension of differences is probably confined to relatively crude and concrete ones, viewed against the familiar background of the child's own social milieu. It is to be expected that this perspective will widen *pari passu* with the multiplication of direct and indirect social contacts which results from attendance at school, with a consequent gradual increase of insight into the range and complexity of social differences.

It was considered probable that the rate of this development would be at least to some extent a function of sheer rise in intellectual capacity. But it was also thought that children from social strata where there is a more-than-ordinary awareness of social status might be specially sensitized to perceive status cues. A phenomenon of this kind is suggested by Allport & Kramer's (1946) studies of prejudice. These authors put forward the generalization, challenged by Carter (1948), but vindicated by Lindzey & Rogolsky (1950), that prejudiced people are better able to identify members of a disliked outgroup than are non-prejudiced people.

II. METHODS

(i) *Subjects*

The study was carried out in Glasgow. In order to get an adequate coverage of the range of socio-economic status among the children's families, five schools were selected: two in a poor working-class district, two attracting mainly middle-class pupils, and one intermediate one.

Theoretically, the best plan for choosing the subjects would have been to ascertain the social background of all the children in the requisite age-groups, and then to draw samples accurately balanced according to sex, age and status variables. Unfortunately this was impracticable, as the information about parental occupations available from the school record cards was insufficient for the purpose. The procedure followed was to list all the children falling into the appropriate age-groups given in years and months: 6.0–6.11, 7.0–7.11, 8.0–8.11, 9.0–9.11; and to use these lists as the sampling frame, from which the subjects were selected by means of a table of random numbers. In each of four schools five boys and five girls at each of the various age levels were selected, totalling forty per school. Later a further school for boys only, and therefore yielding only twenty subjects, was added to make good a deficiency in boys from a middle-class background.

To classify the children, the categories used by Himmelweit, Halsey & Oppenheim (1952) were taken over. These are shown below, together with the abbreviations employed hereinafter:

Upper middle and middle class	UMC	MC
Lower middle class	LMC	
Upper working class	UWC	WC
Lower working class	LWC	

The allocation to these categories was based on three types of information: (i) the above-mentioned school record cards showing parental occupations; (ii) the children themselves were asked about their parents' jobs, and advantage was also taken of such additional clues as were volunteered by the children, e.g. descriptions of their homes, where and how they went on their holidays; (iii) when the interviews were completed, teachers assisted with an assessment of the home-background of each child, usually from their personal knowledge of the parents.

Table 1. *Sex, age and socio-economic status of the sample*

Parental status	Ages 6–	7–	8–	9–	All ages
		(*a*) Boys			
UMC	5	5	8	7	25
LMC	7	8	5	8	28
UWC	4	10	4	3	21
LWC	8	2	9	6	25
All status levels	24	25	26	24	99
		(*b*) Girls			
UMC	3	6	4	2	15
LMC	3	5	4	7	19
UWC	6	5	8	3	22
LWC	7	5	4	8	24
All status levels	19	21	20	20	80

The composition of the sample is shown in Table 1, which requires some comment. One boy and one girl at marginal ages were wrongly assigned originally and thus had to be transferred. One boy had to be eliminated from the sample because the material relating to him was incomplete. It will be seen that the sampling procedure adopted resulted in a distribution among the cells which is not as even as one might have wished. However, this is probably a price worth paying for ensuring the absence of bias, especially as it does not seriously affect the main analysis.

(ii) *The social perception test material, administration and scoring*

In devising a technique for assessing the ability of children to perceive social differences, a central problem was how to enable one to dispense with elaborate verbal instructions. The danger inherent in any such instructions is twofold: first, poor vocabulary might lead to misunderstanding and consequently a spuriously low level of performance; secondly, with extensive instructions it would be hard to avoid giving the children cues about the nature of the task. Hence a pictorial method which minimizes these difficulties was devised.

The underlying *rationale* of the technique rests on the difference between physical and social incongruity. Pictorial tests of intelligence commonly make use of the former, e.g. showing a boat with the smoke from the funnel drifting one way, and the flag being blown the other.

Most of the test material consists of pairs of pictures, each member of the pair being divided into two pieces so as to form a kind of elementary jig-saw. The four pieces can be put together in two different ways, both of which will result in a pair of pictures which are physically correct. There is only one combination, however, that produces a pair of *socially congruent or probable* pictures. The final version of the material, arrived at after extensive preliminary trials, is reproduced in Figs. 1 and 2.

In order to illustrate the principle, Set 1 is shown put together the 'wrong' way; *a–b* and *c–d* would of course be the expected combinations, but the background is so arranged that the more improbable action can be equally well represented. The same applies to Sets 2–4; Set 5 involved a slightly different sorting task which is described later.

The sequence of administration followed the order of the numbering. It was arrived at empirically, on the basis of preliminary trials. Contrary to the normal practice in intelligence tests, the order is one of *decreasing* difficulty. Otherwise a child might grasp the underlying principle from the easy items at the beginning, thus being in a position to make shrewd guesses with the harder problems, which he would not have been able to solve if they had been presented first.

The purpose of the first set, dealing with the neutral subjects of football and cricket, was merely to introduce the subjects to the requirements of the task. With this and the three subsequent sets the cards were placed on the table at different angles in random positions, and the children invited to assemble them so as to make two pictures. This was relatively easy, and nearly all the children were able to put the pieces together correctly at once. When they had done so, and the investigator had made sure the children felt certain of their response, the position of the cards was reversed to *a–d* and *c–b*, as in Fig. 1. The question 'would it be all right like this?' was normally answered in the negative. Younger children were apt to observe that it was 'wrong', whilst older ones would more often remark 'you don't do that', 'it doesn't make sense', or something to that effect. This was then discussed, at any required length, to put over in simple terms the idea that such a combination was neither morally 'wrong' nor physically impossible, but merely unlikely. It was further explained that a series of cards would be shown to which the same idea applied: the pieces could be put together in different ways, and they were always to choose the one that was more usual, to make pictures of things one could often see. The establishment of this 'preparatory set' was of great importance, and considerable care was taken to do it thoroughly.

The mode of administering Sets 2–4 was uniform, and will be explained in conjunction with the method of scoring. As regards the latter, the major problem was the existence of only two alternatives: a picture concordant in terms of social class was the right solution, a discordant picture the wrong one. There was thus one chance in two of being successful by mere guessing. A procedure was devised to get around this difficulty, so that the probability of scoring errors from this source would be minimized, yet without having to rely unduly on verbal reports.

After a child had put the cards together correctly and indicated that the task was completed, the investigator would *reverse* the positions and ask for comments in a neutral manner. If the child was equally satisfied with the resulting socially incongruent picture, or appeared indifferent, the performance was counted a failure; if, on the other hand, the child refused to accept the reversal, a successful performance was recorded. In that case a further inquiry concerning the grounds of rejection followed. This, in addition to the frequent spontaneous remarks at all stages of administration, yielded valuable insights into the children's ideas about social differences. In some doubtful cases other factors were taken into consideration, e.g. whether the child had just pushed the two pairs of cards together and left it at that, or had tried out different positions several times before arriving at a final decision, as many of them did. Where the first response was wrong no problem arose and the failure was clear.

Set 1 served as an introduction, and was not scored. For each of the following three Sets two independent scores were assigned, based respectively on performance (right = 1; wrong = 0) and on verbal response (adequate = 1; inadequate = 0). The former was generally simple and clear-cut, the latter somewhat less easy to score. Reference has already been made to the fact that young children have not yet acquired a distinctive vocabulary of social class at a time when they may be quite capable of making perceptual discriminations. It was therefore thought appropriate to adopt as the criterion of an adequate verbal response whether or not attributes of both people and backgrounds *related to socio-economic status* were mentioned. This is apt to sound rather more vague than it was in practice. As will be seen later, occupational labels were freely employed, or terms connoting degree of wealth; even in the absence of such more or less direct references the criterion usually led to an unambiguous decision. Independent scoring by two judges resulted in 97% agreement on performance, and 92% on verbal response. Nearly all the disagreements turned out to be mistakes in applying the proper criteria, and were thus readily resolved. These reliability figures include Set 5, about which a little more must be said.

Here the procedure was different. The two larger cards with pictures of houses were placed on the table side by side, about 8 in. apart. In the centre and slightly below, all the cards depicting parents and children were stacked in a mixed heap. The children were told that they should put together two families, each consisting of father, mother, boy and girl, and lay out each family on top of the house in which they lived. The task was therefore twofold, namely assembling the families correctly and assigning them to the right house. Hence four scores were given on Set 5, two for performance and two for verbal responses.

For the social perception test as a whole the total range of possible scores was 0–5 on performance and 0–5 on verbal response.

(iii) *Other tests used and general procedure*

Each child was given the Coloured Progressive Matrices and the Crichton Vocabulary Scale. It had also been hoped that the results of the social perception test could be related to another kind of developmental measure, and for a variety of reasons Piaget's (1930) task of explaining the working of a bicycle was chosen. A standardized version was devised, consisting of a cardboard model with movable parts, which the children had to assemble and explain. Although the responses were recorded in every detail, it did not prove possible to allocate the majority of the children unambiguously to any of Piaget's four stages except the last, i.e. where the mechanism is fully understood. Success was mainly confined to the higher age-grades, especially in the case of girls. The original object was thus not achieved, but as will appear shortly, the effort was not altogether wasted.

The procedure followed in each school was as follows: after the sample had been selected, all classes containing prospective subjects had a preliminary visit. On this occasion the whole class was asked to draw a bicycle, the aim being to get some check on later performance; at the same time this meeting served to establish contact with children and teachers.

At a later date the children were called in random order for individual interview. Apart from a single case, there was no indication that the children came to the interview with any notion of what they would be asked to do.

Each interview lasted roughly one hour. After some initial small talk to put the child at ease, during which parental occupation and related matters formed the topic of conversation, a start was made with the progressive matrices. Then came the vocabulary scale, followed by the social perception test. A certain amount of anxiety was occasionally manifest in the course of this, and the subsequent bicycle task provided a suitable and necessary relief from any tension that may have accumulated. This last was greatly enjoyed by nearly all the children.

On the whole *rapport* was excellent, an important factor in an investigation of this nature.

III. QUANTITATIVE RESULTS

(i) *Age changes in the perception of social differences*

The general trends are seen in Fig. 3, where, in view of the small numbers, the social status categories have been reduced to 'middle' and 'working'. As expected, there is a fairly regular age progression in both performance and verbal scores, but the former are consistently higher than the latter.

Age changes may also be described in terms of the incidence of total failure indicated by a score of zero. Some 21% of those aged 6+ received such a score on performance, and 51% on their verbal responses; the corresponding percentages at the age of 9+ were 2 and 7. Thus it would seem that, by the time they are 10 years old, very few children remain totally unaware of social differences.

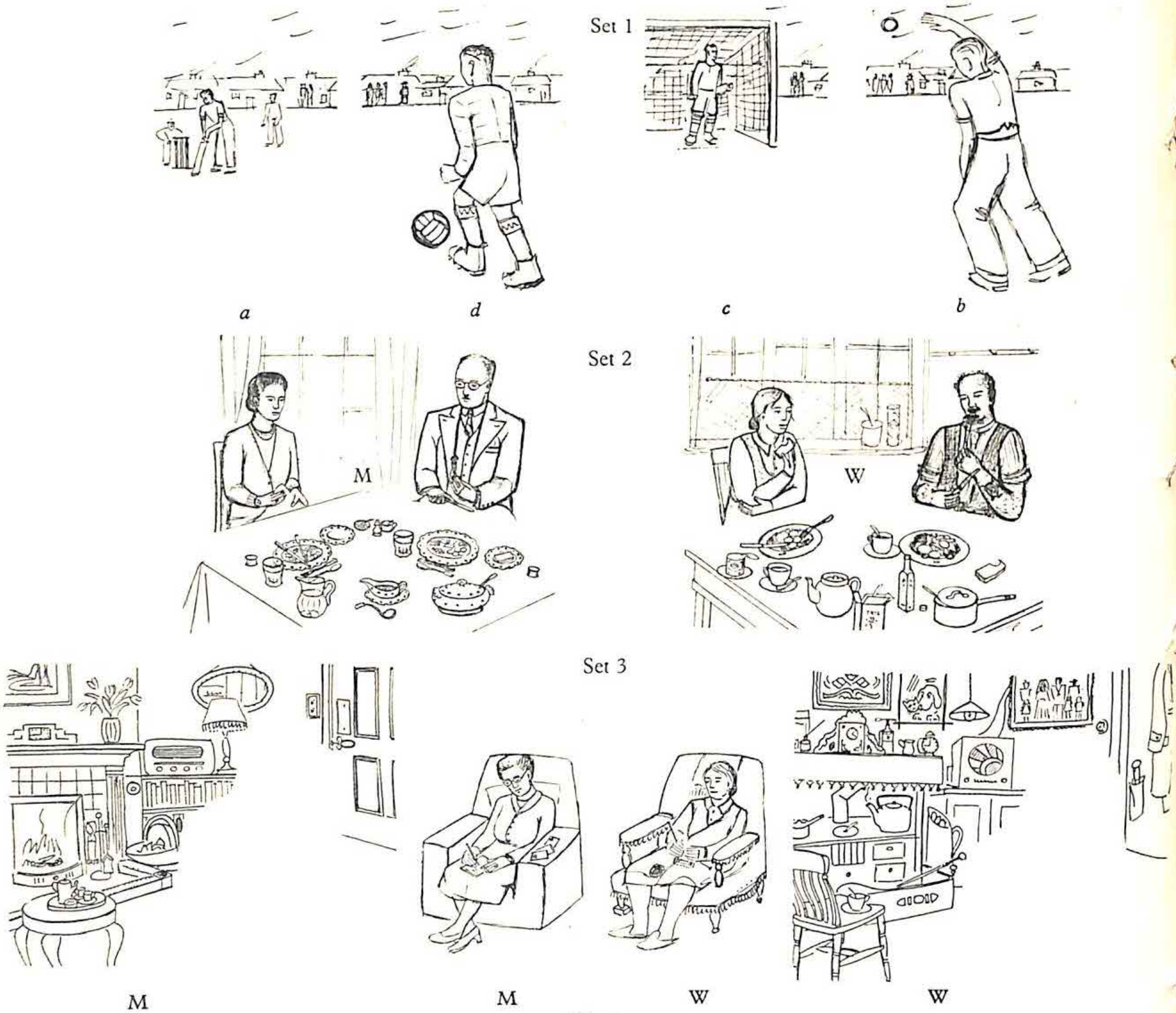


Fig. 1

(ii) *Variations according to sex and social status*

The overall significance of such differences was tested by analysis of variance, following the method given by Lindquist (1940) which makes allowance for unequal numbers in subgroups. From Table 2 it will be apparent that the status differences are highly significant as regards both performance and verbal scores, while the interaction reaches significance only with the former. The sex difference as such is negligible.

The scores of the various subgroups, set out in Table 3, can now be considered in some more detail. As far as status is concerned, the major dividing line is the one separating UWC from LMC; three of the critical ratios reach significance, and the remaining one closely approaches it. Differences between UMC and LMC on the one hand, and UWC and LWC on the other are relatively small.

The sex-status interaction on performance is reflected in the fact that MC boys have lower, and WC boys higher, scores than girls at the corresponding status levels. The critical ratios of sex differences were not computed in the case of verbal scores, as the results of the analysis of variance do not warrant it.

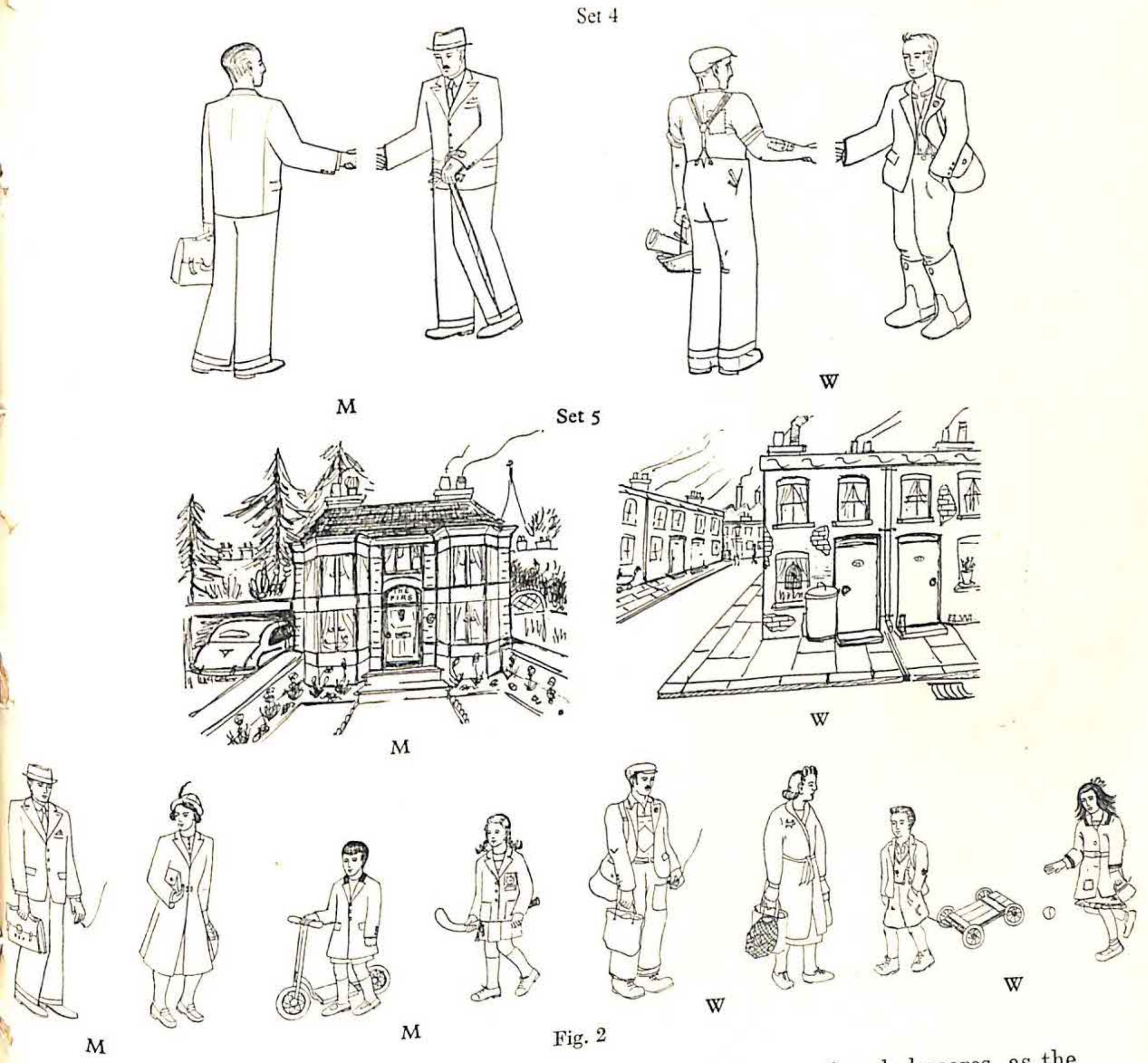


Fig. 2

What emerges most clearly from the table is the special position of the LMC girls. In absolute terms their mean performance and verbal scores are the highest among all the subgroups, and they account for all the differences significant at the 0·001 level.

(iii) *Intelligence as a factor in social perception scores*

What has been reported so far establishes the existence of some sex and status differences, but these are open to at least two main interpretations: first, they may be attributed entirely to the well-documented association (e.g. Burt, 1943) between intelligence and social status; secondly, it may be thought that quite apart from intelligence, contrasting social environments are likely to affect children's disposition to perceive social differences.

The data on intelligence, shown in Table 4, will serve to resolve this question. A few comments on this table are necessary. There is the expected decline in mean test scores

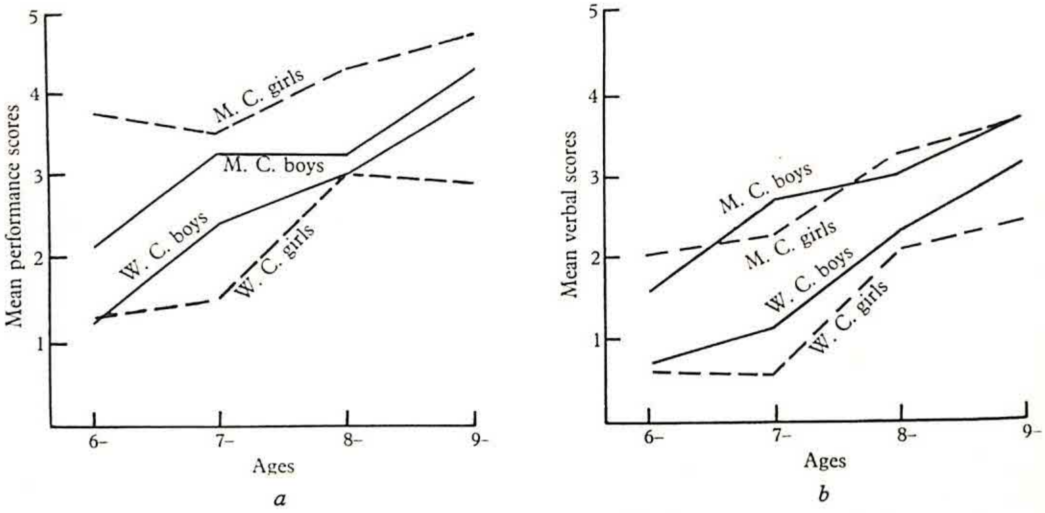


Fig. 3. Age-changes in mean social perception scores. (*a*) Performance scores. (*b*) Verbal scores

Table 2. *Analysis of variance of social perception test scores*

		Performance			Verbal		
Source	D.F.	Mean square	*F*	*P*	Mean square	*F*	*P*
Between groups:							
Sex	1	1·01	0·49	—	1·19	0·46	—
Status	3	24·85	12·05	< 0·001	21·45	8·35	< 0·001
Sex × status	3	5·74	2·78	< 0·05	0·82	0·32	—
Within groups	171	2·06	—	—	2·57	—	—

Table 3. *Sex and status distribution of Social Perception scores*

(Critical ratios given in parentheses.)

Parental status		Performance			Verbal	
		Boys		Girls	Boys	Girls
UMC	Mean	3·28	(1·27)	3·80	2·92	2·67
	S.D.	1·22		1·33	1·26	1·74
		(0·30)		(1·56)	(0·51)	(0·47)
LMC	Mean	3·39	(3·30)*	4·37	2·71	2·95
	S.D.	1·54		0·58	1·66	1·70
		(1·85)		(5·66)*	(2·23)†	(3·16)*
UWC	Mean	2·67	(0·91)	2·41	1·71	1·41
	S.D.	1·21		1·61	1·45	1·37
		(0·25)		(0·87)	(0·18)	(0·01)
LWC	Mean	2·56	(1·21)	2·00	1·80	1·46
	S.D.	1·65		1·58	1·72	1·61

* *P* less than 0·001.

† *P* less than 0·05.

with decreasing status; one exception is the UMC girls, who scored slightly lower than the LMC ones on the matrices. A surprising feature is the uniform superiority of the boys on both tests, which could hardly be attributed to the sampling. However, as the present object is to assess the influence of intelligence separately *within* each sex category, this may be safely ignored.

An analysis of co-variance was carried out, in which the social perception test scores are adjusted so as to make allowance for differences in intelligence. When this is done, as set out in Table 5, the general picture is substantially modified. The verbal scores, adjusted

Table 4. *Coloured Progressive Matrices and Crichton Vocabulary scores*

Parental status	Sex	Matrices Mean	Matrices S.D.	Vocabulary Mean	Vocabulary S.D.
UMC	B	26·44	4·62	48·20	11·14
	G	21·73	5·21	44·13	10·46
LMC	B	23·07	5·21	43·43	9·94
	G	22·63	6·43	41·42	12·00
UWC	B	19·71	3·93	32·24	10·33
	G	18·05	4·99	31·09	12·50
LWC	B	17·64	5·09	27·48	13·54
	G	15·21	3·81	25·17	7·92

Table 5. *Analyses of covariance of social perception scores: variance ratios*

	Performance scores adjusted for matrices		Verbal scores adjusted for vocabulary	
	Boys	Girls	Boys	Girls
F	1·21	6·17	0·65	1·65
P	—	< 0·001	—	—

Table 6. *Correlations between Pintner General Ability Test scores and Stendler's Picture Test scores (Brasstown) compared with those between Matrices scores and Social Perception Performance scores (Glasgow)*

	Brasstown			Glasgow		
Ages	*N*	*r**	*P*	*N*	*r*	*P*
6 and 7	29	− 0.09	—	89	0·40	< 0·001
8 and 9	24	0·47	< 0·02	90	0·39	< 0·001

* The signs of the Brasstown *r*'s were reversed so as to make them correspond in meaning with the Scottish ones.

for vocabulary as assessed by the Crichton Scale, no longer show any significant status differences with both sexes. The boys' performance scores at the various status levels are also wholly accounted for in terms of the relationship between these scores and intelligence, as measured by the Coloured Progressive Matrices. Only the performance scores of the girls continue to show highly significant status differences after the adjustment for intelligence has been made. It follows that intelligence is largely, though not wholly, responsible for the results obtained.

Lastly, the material permits a restricted comparison with Stendler's findings. Stendler (1949) correlated her subjects' intelligence and picture test scores, and in Table 6 her coefficients are juxtaposed with the corresponding ones from the present investigation. In this connexion Stendler (1949, p. 56) commented as follows: 'The almost-zero correlation on the first-grade level [i.e. ages 6 and 7] may be another indication of the fact that most children of this age are not yet dealing with the problem of social class.... Apparently the problem has little meaning for [the young child].' But with similar age-groups the Scottish coefficients came out as highly significant. While the possibility of a national difference cannot be excluded, it appears unlikely in view of the parallel research findings in Britain and U.S.A. for adolescents and adults. These considerations suggest that the reason for the near-zero correlation is to be sought in the particular way in which the problem of social class was presented to the young Brasstown children.

IV. LANGUAGE AND EMERGING CONCEPTIONS OF SOCIAL CLASS

(i) *Nature of the task and types of failure*

For an adequate appreciation of the factors that determine the responses to the picture, a brief general outline of the skills involved is necessary. The task was basically one of matching representations of material settings with those of people whose status could be recognized from the cues provided, and of similarly matching people of the same social status. In order to be able to accomplish this, two things must have been learnt previously by direct or indirect experience: first, the visual cues whereby people of contrasting social status can be distinguished, including their characteristic material environment; and secondly, the behavioural norms governing the relationships between people at different status levels. A further requirement is that the pictures as a whole should be interpreted correctly, i.e. in the sense intended by the investigator.

Success in the subsidiary verbal task presupposes the acquisition of a vocabulary relating to social differences, or at least sufficient verbal skill for appropriate indirect references to such differences.

A wider aspect, implicit in what has just been said, concerns the formation of concepts relating to social class. The child must be able to abstract from individual elements and to use the patterned wholes they constitute in combination to assign these wholes to a broader category. Such a concept may be present, in the sense that certain parts may be felt to belong together, before a suitable verbal label can be attached to the category. This was demonstrated by Hanfmann (1941) in a perceptual situation involving the sorting of objects.

The main kinds of failure may be discussed in the light of these considerations. First there were those among the younger or duller children who clearly felt completely at sea when confronted with the pictures; in spite of the practice given with the introductory set, they tried to deal with the pieces on the basis of their experience with jig-saw puzzles, looking for optimum physical fit. When this proved fruitless, they tended to seize on some irrelevant details, e.g. pairing respectively those with and without a hat. Such children were as a rule quite content with the reversal, both combinations being for them equally devoid of any special significance.

Such children failed throughout the test and scored nil—one might regard them as being in the 'no class concept' stage. Older or brighter children often responded similarly, but only to the earlier and more difficult items. Certain cues did not seem to be included in their repertoire and thus could not call forth the general class concept, which in any case probably had a rather low saliency. One of the older boys, for example, whose comment is quoted below, placed the working-class couple at the middle-class table.

(In this and all subsequent quotations [W] refers to working-class and [M] to middle-class figures or backgrounds in the stimulus pictures.)

> [Reversal] No; this lady [W] is eating bread, but she hasn't got a plate for eating bread off. (UMC boy, 9, Set 2)

This boy was attempting to make the picture suit his own middle-class norms of table etiquette, quite oblivious of the glaring social incongruity which resulted therefrom. Errors sometimes also occurred where a child was aware of a social difference, but lacked information about the environmental concomitants of social stratification:

> These people [W] look poorer than these people [M] and they [W] have no teapot and no sauce. (LWC girl, 9, Set 2)

This girl made the same mistake for an entirely different reason, feeling that a table where the teapot and bottle of sauce were missing could not be that of the wealthier people.

Only one instance was encountered where apparent ignorance of the norms governing social relationships between people of different status was coupled with adequate recognition of social differences. It is likely that both are normally learnt at the same time.

In a few cases difficulties arose when figures supposed to represent working-class people were taken to be farmers, as in our society these have no clear-cut position within the status hierarchy. Similarly an urban-rural dichotomy was occasionally substituted for the social class one. Here is an example of a boy who put the working-class family into the villa, and the middle-class one into the tenement, with the following remarks:

> That [M] is a house on the edge of the country and they don't have such good shops to buy clothes. They [M] are in the town and they have got plenty of good shops to buy good clothes. (UMC boy, 7, Set 5)

Turning now to verbal failures, it has already been said that these were not incompatible with a well-developed covert concept of class. Children were frequently quite emphatic about which was the right and which the wrong way, without being able to explain their preference. Some merely reiterated that the reversal was wrong; more commonly an effort was made to account for one's choice by some kind of descriptive reference to the content of the picture. Both types are illustrated in the following protocols:

> [Reversal] No. [Why?] Because they are the wrong way round. [What do you mean?] They are at the wrong person. [How do you know?] Because they are not the right one. (UMC girl, 7, Set 4)

> [Reversal] That's silly. [Why?] Because...[puzzled pause]...they [W] are in the kitchen.... (UWC boy, 6, Set 2)

Leaving out the cases of total failure, the age changes in modes of response may now be roughly outlined. The youngest children relied chiefly on global impressions, which they were as a rule unable to break down for the purpose of giving an explanation. At an intermediate stage such global impressions began to be translated into at least partially evaluative judgements, in the sense of not being closely tied to the information contained in the stimulus pictures; examples would be comments like 'they are untidy' or 'these look more beautiful'. Actual details mentioned at this stage were usually confined to such highly visible ones as type or quality of clothing. Older children, whilst continuing to bring in evaluative judgements, were often able to pick out relatively inconspicuous signs indicative of social differences; for instance, they might point to the cigarette perched on the ear of one of the [W] figures, or mention the school uniform worn by the little girl in the [M] family. Thus there was evidence of a gradual learning of specific signs which enabled the child both to make increasingly accurate discriminations, and to verbalize the cues on which they were based. Nothing has been said about the expressions employed in referring to social differences, for this topic deserves special treatment.

(ii) *The vocabulary of social class*

At the outset it is important to note that the word 'class' itself was never uttered by any of the children. This is hardly surprising, in view of the fact mentioned earlier that a majority of 13-to-14-year-olds were found to be unfamiliar with the term. The children therefore had to characterize social distinctions more or less indirectly, and the distribution of the various ways in which this was done is set out in Table 7. The numbers stand for the proportions of *different* words used by children in the respective categories; i.e. a word occurring in several responses of the *same* child was counted as only one unit.

Table 7. *Types of class-oriented descriptions of people*

(Rounded percentages based on combined number of references by boys and girls.)

	Categories of children			
	Working class		Middle class	
Descriptions relating to	6 and 7	8 and 9	6 and 7	8 and 9
Working-class figures ...	$N = 12$	$N = 36$	$N = 39$	$N = 58$
(1) Workers	67	47	49	38
(2) Occupational	25	6	18	14
(3) Poor	—	39	28	45
(4) Other	8	8	5	3
Middle-class figures ...	$N = 9$	$N = 26$	$N = 21$	$N = 50$
(1) Gentleman	22	8	5	8
(2) Occupational	56	35	29	26
(3) Rich	—	54	19	38
(4) Other	22	4	47	28

It is necessary to enlarge somewhat upon the actual words and phrases covered by the headings. Starting with those relating to working-class figures, no. 1 was restricted to the terms 'worker', 'workman' or 'working man'; no. 2 included manual jobs, mostly skilled, such as painter, plumber, joiner or electrician; no. 3, in addition to the actual word 'poor' which was most frequent, also covered expressions like 'has not much money' or 'is not well off'; the residual and relatively unimportant no. 4 consisted either of negatively phrased references ('not respectable') or of statements denoting an employer-employee relationship ('this man was going to work for that man').

Turning now to the middle-class figures, 'gentleman' was the only word admitted under no. 1; the occupations mentioned in connexion with no. 2 were, in descending order of frequency: 'in an office', 'business man', 'doctor' and a few miscellaneous ones; no. 3 included the term 'rich' itself and several near-synonyms like 'has lots of money' or 'can afford more'; no. 4 contained a multitude of oblique class references, in which the word 'nice' featured prominently, e.g. 'these are *nice* people', the adjective being usually stressed in actual speech; other expressions were 'good', 'grand', 'posh' and 'important'.

Several trends may be distinguished within this pattern of usage. All categories of children seemed to find it easier to characterize the working than the middle class, as indicated by the uniformly smaller vocabulary relating to the latter. It may be noted that it was not really necessary to attach a verbal label to both in order to explain one's solution: many children said, for instance, 'he is a worker and he is not'. The term 'gentleman' was applied only to middle-class figures, whilst middle- and working-class ones alike were given the title of 'lady'.

A striking fact was that none of the younger WC children employed the words 'rich' or 'poor', yet nearly half the older ones did so; a similar, though less pronounced tendency was found among the MC children ($P < 0{\cdot}001$ and $0{\cdot}05$ respectively). Furthermore, in talking about middle-class figures the MC children were significantly more apt to resort to indirect expressions of the kind described ($P < 0{\cdot}01$).

In the process of abstraction and categorization entailed by the transformation of the data into numerical form, much is lost of the subtler aspects of language. Here are two examples:

> They [W] have a wooden table and they [M] have a nice table. They [W] look as though they would have a wooden table. (LMC girl, 6, Set 2)
>
> It doesn't look right. A well-dressed lady in a kitchen writing letters. And a lady in an old worn chair in a lovely room—it isn't right. (UWC girl, 9, Set 3)

Even when the children did use a particular word, its connotations were subject to considerable variation, and this becomes evident only from the context. This was particularly true of terms related to work, where there was a frequent groping for distinctions that were but dimly discerned.

> He's not working. [What do you mean?] 'Cause he's a doctor. (LWC boy, 6, Set 4)
>
> They [W] are working men. Maybe they [M] are working men too, but they work in a different work. (LWC boy, 9, Set 4)
>
> A workman [W]...and another kind of workman [M]. [What kind?] Well, a shopkeeper or something. [His father is one.] (LMC boy, 6, Set 4)
>
> I think these two are workers [W] and these are just working *men* [M]. (UMC boy, 9, Set 4)

These illustrations show why Table 7 should be interpreted with caution. It is constructed on the assumption that formally identical words used by children at different ages and from different status levels can be treated as equivalent, but this is clearly an over-simplification. The material is not extensive enough, and the shades of meaning too elusive, for confident generalization. However, it may be worth recording a mere impression: it seems that younger children employ occupational terms mainly in a functional and descriptive sense; with increasing age it seems that such terms are becoming more value-loaded—apparently the children are beginning to realize that people doing different jobs vary also in other ways; in short, an awareness of the social status hierarchy is gradually crystallizing.

(iii) *Other aspects of the verbal responses*

It is widely held that social class relationships involve feelings of inferiority, superiority and equality. In the present study these rarely emerged in a clear-cut fashion, possibly due to the fact that only Set 4 presented the issue sufficiently sharply. The few responses of this type were made by older children:

> These are two important men and these are just workmen.... (UMC boy, 8, Set 4)
>
> He's rich, and he's rich, and they [W] are in their working clothes; and they [W] are not as good as them. (LWC boy, 9, Set 4)

On the other hand, it was common for the MC children to take their own social level as the frame of reference and talk about working-class people in terms of a deviation from this norm; and the phrasing often implied a certain amount of condescension:

> It isn't usual for a person to shake hands with a painter really. (UMC girl, 8, Set 4)

They also frequently used the terms 'ordinary man' or 'ordinary business man' to refer to the middle-class figures. On the other hand, none of the WC children suggested in any way that the normal thing is to be a worker. Two, one of whom is quoted below, implied that skilled men are better than other people. A majority of the older children merely stated in a matter-of-fact way that people of different kinds would have nothing to do with each other.

> He is a joiner, but he [M] is only an ordinary man. (UWC boy, 7, Set 4)
>
> This is a worker and this isn't. If that [M] was that man's boss, why would he want to shake hands with him? (LWC boy, 9, Set 4)

Many children felt impelled to account for the fact that [W] men do not as a rule shake hands with [M] men. Older boys tended to put forward the reasonable view that people doing different work were unlikely to meet, and thus could not get to know each other:

> Usually people not so wealthy know each other—they work together. (UMC boy, 9, Set 4)
>
> That man doesn't work beside him, so he doesn't know him. (LMC boy, 8, Set 4)

The ideas of a few of the younger children were very naïve, and most of the girls, irrespective of status, spoke about dirty hands.

> You don't usually see a man shaking hands with a builder, because he's usually hard at work.... (UMC boy, 7, Set 4)
>
> You wouldn't like to shake hands with a workman, because their hands are dirty; it's much nicer the other way. (UMC girl, 8, Set 4)
>
> If that man [M] shook hands with that man [W] his hands would get all paint. (LWC girl, 7, Set 4)

There was in reality no direct indication in the pictures that the workmen had dirty hands. It would not have been practicable to contrive this, as the joint cut across two clasped hands and both combinations had to be equally feasible. It is perhaps no accident that the workers were often described as painters, for this makes such an interpretation more plausible.

A tendency shared by all children was that of applying epithets like 'messy', 'dirty' or 'untidy' to the [W] figures on the pictures and their environment; other terms were used in this way only by MC children, e.g. 'scruffy', 'slummy', 'ugly', 'not well-mannered'. Here are examples of comments hardly justified by the actual content of the stimulus pictures:

> This lady [M] looks as though she would get up and clear cups off chairs and things. This lady [W] looks quite lazy.... (LMC girl, 9, Set 3)
>
> They [W] all look horrible—probably because they live in that house. (UMC boy, 9, Set 5)

The emotional colouring of the children's responses was often manifest mainly in their tone of voice, and is therefore difficult to convey. It ranged from cool equanimity to shocked rejection. Most commonly the responses suggested that the situation represented by the reversal was felt to be unfitting or absurd. Expressions used included 'it would be funny', 'it looks daft', 'it's not *nice* for a man to...' and 'this doesn't look natural'. The following illustrations are drawn from the extremes of the range:

> I would say that was quite all right, but you don't usually see these kinds of men shaking hands. (UMC boy, 9, Set 4)

> ...it wouldn't be right—it would be awful for these people [M] to live there. (UWC girl, 9, Set 5)

Outright emotional reactions did not occur among the younger children. WC children rarely showed any embarrassment when discussing social differences, but MC ones not infrequently did—a few quite acutely. WC children used the words 'rich' or 'poor' without qualification; MC ones had some inhibitions about these words, and unless they employed circumlocutions there was apt to be some hesitation:

> Usually the tattooed people are not so [Pause]...well, they are poorer. (UMC boy, 9, Set 2)

Finally it may be noted that in seventeen cases the differences between the people represented were interpreted in terms of *age*. This trend was confined to WC and LMC children, and some explained their ideas quite clearly:

> The younger people [M] work harder. They can afford nice things. The older people [W] wouldn't have nice dishes like that. (LMC girl, 7, Set 2)

> These people [W] look older and these people [M] look younger. They [M] would be working, so they can get good things. (LWC boy, 8, Set 2)

In concluding this brief survey of some of the qualitative aspects of the responses a reminder is indicated that these were either spontaneously offered by the children, or elicited by strictly neutral probing. Whilst this provides a guarantee that the responses accurately reflect the children's ideas and were not suggested to them by the investigators, the absence of direct control has the drawback that it leads to lack of uniformity. While the resulting generalizations are inevitably somewhat vague and impressionistic, it is believed that such an approach is essential for an understanding of the development of children's ideas and attitudes.

In the earlier part of the work it has been established that, for nearly all categories of children, the ability to perceive social differences is largely a function of intellectual growth. The material presented in the last three sections strongly suggests that attitudes and feelings about social differences, as distinct from their sheer apprehension, are influenced to a much greater degree by the specific social background of each child.

V. DISCUSSION

In assessing the implications of the findings, it has to be remembered that the children were faced only with certain limited facets of class differences, such as could be incorporated into pictorial material. Furthermore, as regards the representations of working-class people, some slight exaggerations were necessary so as to avoid making the task too difficult. About a generation ago the sartorial contrast between middle- and working-class people, and their children, was a good deal greater than it is now. In Set 5 the working-class children were shown in rather more tattered outfits than those worn at school by a majority of the working-class respondents. It was therefore less likely that these children would wholly identify with their pictorial counterparts than the middle-class children. On the other hand working-class housing conditions in Glasgow tend to be poor, and the representations were not far off the mark as far as that was concerned. Nevertheless, it is possible that there was a difference in the impact of the pictures according to the respondents' social background: middle-class children saw their own environment, including people, fairly accurately mirrored, while some of the working-class children, especially children of skilled men, may have felt themselves to be somewhat better off than the working-class figures on the pictures. One cannot be sure about the precise effects of this on the verbal responses. It is likely that it was responsible for a certain quality of detachment discernible in some of the comments by working-class children; and perhaps also the rather liberal sprinkling of adjectives like 'dirty' and 'poor', which they might not have applied if they had viewed the setting as a direct representation of their own.

One is also tempted to speculate about the tendency to call working-class figures on the pictures 'old'. There is a need for caution in this connexion: in real life working-class adults, in particular women of the older generation, are liable to show their years more than a middle-class person of the same chronological age. If this was the case with the stimulus pictures, the answer would be simply that some children with a limited awareness of social class had singled out an irrelevant, though genuine variable. In order to test this possibility, the pictures were submitted to a dozen adults of varying social status with a request to estimate the ages of the people depicted. There was in fact no consistent overestimation of the ages of the working-class figures as compared with the middle-class ones, so it would seem that the above explanation does not hold. Moreover, it may be inferred from the actual phrasing used by several of the children that their age judgements are based not just on the appearance of the person, but on the person-in-his-setting. This suggests the following interpretation: working-class children are familiar with the life-cycle of working-class families, in which retirement often means living on a very reduced income. Seeing people in poor surroundings, therefore, may lead them to suppose that they are old. On the other hand, lower middle-class children who responded in this way may have come from families that were socially mobile; and the pictures of people in a materially less favourable environment than their own may remind them of their grandparents. Naturally this is merely a tentative hypothesis, and the numbers are too small for any attempt at verification from the present sample.

Perhaps the most striking aspect of the verbal responses is the diversity of notions related to social differences which they reveal. Clearly a multitude of influences have been at work to produce a variety of attitudes and perspectives. Whilst they do overlap in certain important areas, there is rather less agreement than one would expect from perusal of the general literature in this field. One could try and account for the discrepancy by the assumption that increasing age exposes the individual to social pressures making for greater uniformity. However, the results as they stand lend support to the contention by Bott (1957) that there is no one valid way of finding out what people really think about social class.

The writer is indebted to Miss S. Clutton who prepared the drawings, and to Miss M. M. Duncan and Mr A. D. Chalmers for their assistance with the field-work. Thanks are

also due to Dr H. S. Mackintosh, Director of Education, who kindly provided the necessary facilities for the research, and to the University Court for supporting it with a grant.

REFERENCES

ALLPORT, G. W. & KRAMER, B. M. (1946). Some roots of prejudice. *J. Psychol.* **22**, 9–39.

BOTT, E. (1957). *Family and Social Network.* London: Tavistock Publications.

BURT, C. (1943). Ability and income. *Brit. J. Educ. Psychol.* **13**, 83–98.

CARTER, L. F. (1948). The identification of 'racial' membership. *J. Abnorm. (Soc.) Psychol.* **43**, 37–53.

CENTERS, R. (1950). Children of the new deal. *Int. J. Opin. Attit. Res.* **4**, 315–36.

GINSBERG, M. (1930). Class consciousness. *Encyclopaedia Soc. Sci.* **3**. New York: Macmillan.

HANFMANN, E. (1941). A study of personal patterns in an intellectual performance. *Character & Pers.* **9**, 315–25.

HIMMELWEIT, H. T., HALSEY, A. H. & OPPENHEIM, A. N. (1952). The views of adolescents on some aspects of the social class structure. *Brit. J. Sociol.* **3**, 148–72.

JAHODA, G. (1953). Social class attitudes and levels of occupational aspiration in secondary modern school leavers. *Brit. J. Psychol.* **44**, 95–107.

LINDQUIST, E. F. (1940). *Statistical Analysis in Educational Research.* Boston: Houghton Mifflin.

LINDZEY, G. & ROGOLSKY, S. (1950). Prejudice and identification of minority group membership. *J. Abnorm. (Soc.) Psychol.* **45**, 37–53.

MACE, C. A. (1938). Beliefs and attitudes in class relations. In T. H. Marshall (ed.): *Class Conflict and Social Stratification.* London: Le Play House Press.

MACRAE, D. G. (1953). Social Stratification. A trend report and bibliography. *Current Sociology*, **2**. Paris: UNESCO.

MARTIN, F. M. (1954). Some subjective aspects of social stratification. In D. V. Glass (ed.): *Social Mobility in Britain.* London: Routledge and Kegan Paul.

PFAUTZ, H. W. (1953). The current literature on social stratification: critique and bibliography. *Amer. J. Sociol.* **58**, 391–418.

PIAGET, J. (1930). *The Child's Conception of Physical Causality.* London: Routledge and Kegan Paul.

STENDLER, C. B. (1949). *Children of Brasstown. Their Awareness of the Symbols of Social Class.* Urbana: University of Illinois Press.

(*Manuscript received* 10 *March* 1958.)

THE DIFFERENTIATION BETWEEN NORMAL AND VARIOUS NEUROTIC GROUPS ON THE MAUDSLEY PERSONALITY INVENTORY

By H. J. EYSENCK

University of London

The M.P.I. (Eysenck, 1958*b*) is a 48-question inventory which purports to measure, in a rough and ready fashion, the two personality dimensions of Neuroticism (N) and Extraversion-Introversion (E). It has been carefully constructed on the basis of item- and factor-analyses (Eysenck, 1956), possesses adequate reliability (Jensen, 1958), and has been shown to possess 'construct validity' (Eysenck, 1957); the two scores are independent of each other. A summary of researches on a variety of normal groups (students, nurses, industrial apprentices, etc.) as well as several neurotic groups has shown considerable differentiating power as far as the N scale is concerned (Jensen, 1958); a study of the responses of different neurotic groups (Sigal, Star & Franks, 1958) found psychopaths and hysterics more extraverted than dysthymics, but did not succeed in obtaining statistically significant results due to the small numbers employed (Eysenck, 1958*c*). The present note is intended to collect together all available information on the differentiation between various neurotic groups, between normals and neurotics, and between all the above groups and recidivist criminals and patients suffering from psychosomatic disorders.

The means and S.D.'s are given in Table 1. The normal standardization group consisted of a quota sample of the population, collected along similar lines to those described elsewhere for the standardization of the short form of the M.P.I. (Eysenck, 1958*a*), except that subjects filled in the questionnaire instead of E. The American student group was largely collected by Dr A. W. Bendig of the University of Pittsburgh. The dysthymic, hysteric and psychopathic groups are made up of hospital patients so diagnosed by experienced and responsible psychiatrists, or else unanimously diagnosed on the basis of their case papers by three independent psychologists (Sigal *et al.* 1958). The prisoner group consisted of recidivists in one of H.M. prisons. The psychosomatic group consisted of patients suffering from asthma and ulcers; their use was suggested to me by Dr A. Barendregt who, in an unpublished paper read at the 1958 International Congress of Applied Psychology, found such groups to be highly neurotic and less introverted than dysthymics on the Heron questionnaire (Heron, 1956), which is in certain ways similar to the M.P.I. and correlates with it.

Table 1

Description of group:	Size	N mean	N_σ	E mean	E_σ
Normals (English): quota sample	1800	19·89	11·02	24·91	9·71
Normals (American students)	1500	20·91	10·69	28·53	8·28
Dysthymics (hospital patients)	84	38·18	10·84	17·86	10·02
Prisoners (recidivists)	146	30·35	10·73	24·09	9·11
Hysterics (hospital patients)	58	30·82	11·84	24·91	9·26
Psychosomatics (hospital patients)	108	35·69	10·89	25·38	9·33
Psychopaths (hospital patients)	36	35·58	10·91	30·77	9·51

The results as tabulated in Table 1 show that all neurotic, psychosomatic and recidivist groups are more neurotic than is the normal sample, that dysthymics are the most introverted and psychopaths the most extraverted groups, and that hysterics, recidivists, and psychosomatic patients cannot be discriminated from each other on the basis of the M.P.I., although they are differentiated from dysthymics, normals and psychopaths. On the whole, these results are in line with expectation (Eysenck, 1957), although the failure of the hysterics to be more extraverted than the normals is somewhat unexpected (Eysenck, 1947). This finding may be connected with the fact that the regression line of N on E is significantly bent at the extreme introverted end, showing a marked tendency, both in normal and in neurotic groups, for subjects with very low extraversion scores to have unduly high neuroticism scores (Eysenck, 1958*b*). On the whole, the results suggest the possible usefulness of this questionnaire in clinical, experimental and applied studies.

REFERENCES

EYSENCK, H. J. (1947). *Dimensions of Personality.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1956). The questionnaire measurement of neuroticism and extraversion. *Riv. Psicol.* **50**, 113–40.

EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1958*a*). A short questionnaire for the measurement of two dimensions of personality. *J. Appl. Psychol.* **42**, 14–17.

EYSENCK, H. J. (1958*b*). *Manual of the Maudsley Personality Inventory.* London: University of London Press.

EYSENCK, H. J. (1958*c*). Hysterics and dysthymics as criterion groups in the study of introversion-extraversion: a reply. *J. Abnorm. (Soc.) Psychol.* **57**, 250–2.

HERON, A. (1956). A two-part personality measure for use as a research criterion. *Brit. J. Psychol.* **47**, 243–51.

JENSEN, A. (1958). The Maudsley Personality Inventory. *Acta psychol., Hague,* **14**, 314–25.

SIGAL, J., STAR, K. H. & FRANKS, C. M. (1958). Hysterics and dysthymics as criterion groups in the study of introversion-extraversion. *J. Abnorm. (Soc.) Psychol.* **57**, 143–8.

(*Manuscript received* 28 *June* 1958)

PUBLICATIONS RECENTLY RECEIVED

The Nature of Experience. By Sir Russell Brain. London: Oxford University Press. 1959. Pp. 73. 8*s.* 6*d.*

Brain—Memory—Learning: a Neurologist's View. By W. Ritchie Russell. London: Oxford University Press. 1959. Pp. xii + 140. 18*s.*

Two eminent neurologists have here written two very different books. Sir Russell Brain's evident breadth of interest made him an excellent choice as Riddell Lecturer at the University of Durham, and awareness of the pitfalls awaiting the amateur philosopher has not prevented him from giving a spirited account of his views on the problems of perception. He believes that philosophers who discuss such abnormalities of perception as hallucinations and illusions may not always know as much as is desirable about their empirical features, and devotes most of his first lecture to providing such information. He then proceeds to argue in favour of a sense-datum theory of perception, and from a discussion of subjectivity as exhibited in the psychoanalytic mechanism of projection he passes in his final talk to the differences between artistic and scientific experience. His book is not addressed to the empirical psychologist, who may nevertheless pick up a point of interest here and there.

Dr Ritchie Russell, on the other hand, addresses himself to the psychologist far less than the title of his book might suggest; he appears, in fact, to take a somewhat archaic view of the subject. This is not only a disappointment to the psychological reader, but also perhaps a loss to the book itself. The central chapters deal with concussion and the traumatic amnesias, epilepsy, the phantom limb, pain and the fronto-hypothalamic system. While it may be hoped that neurological observations of these phenomena will throw some light on the physiological mechanisms of learning and memory, much of the evidence is behavioural in nature, and so might perhaps most readily have been discussed in the light of behavioural concepts and models. The author has chosen instead to rely on a physiological model whose properties are not too clear. The model rests essentially on the notions of spontaneous neural activity and structural change at the synapse, and its explanatory power is weak, so that its application to the neurological material is not as illuminating as could be wished. The learning theorist who is interested in this material (as he should be) will therefore have to do much of the work for himself, and it must be said that he will not be helped by the author's presentation. The book is intended to be intelligible to readers with limited physiological knowledge, but is not so successful in this aim as it might have been. At the behavioural level, on the other hand, the terminology is too loose; a more discriminating use of terms like retention, recognition, recall and remembering might have been particularly helpful. Parts of the book have been developed from previous publications and this may, to some extent, account for its diffuse and repetitive style.

N. E. LOVELESS

Psychopathology: a Source Book. Ed. by C. F. Reed, I. E. Alexander and S. S. Tomkins. Harvard University Press. London: Oxford University Press. 1958. Pp. xii + 803. No price given.

This book is likely to trap the unwary browser. It is impressively covered, well produced and the list of authors is studded with 'star' names in the field of psychopathology. It thus seems at first glance to be a concerted attempt by the best authors to cover areas in psychiatry and clinical psychology. However, it is only when he reads the unobtrusive preface that the reader learns that this is a selection of papers which have already appeared in recent psychological and psychiatric journals.

This may disappoint the clinical or medical psychologist who is likely to have read the papers already. To be fair, the selection is not intended for him but for undergraduate students of abnormal psychology. The preface indicates that this collection of papers is intended to supplement text-books ordinarily used in undergraduate teaching. If this stated intention is valid then the American undergraduate in psychology is a highly sophisticated person. It is doubtful if the undergraduate student of psychology in this country is likely to benefit from all of the papers, particularly those giving accounts of specific experiments in biochemistry, for example.

Scientific papers are necessarily complete in themselves and accordingly do not make good chapters in a book unless they are given continuity and consistency by active editing and commentary. Unfortunately nothing of this kind has been done. Although the instigators describe themselves as editors, their task has been simply to select and arrange papers which have previously been individually edited.

Of the 46 papers, 10 are classified under the heading of 'Psychopathology and Early Experience', 11 under 'Psychosomatic Disorders and Neurosis', 9 under 'Schizophrenic Psychoses', 8 under 'Somatic Factors in Psychopathology' and 8 under 'Psychopathology and the Social Context'. Gaps are to be expected and occur here by the omission of papers on the affective psychoses, psychotherapy, genetics and the newer neurophysiological concepts. While each single contribution is excellent, the fact that it is autonomous exaggerates the uneven quality of the book as a whole. The section on somatic factors, for example, contains the most recent of Klebanoff's encyclopaedic reviews of the Psychopathological Consequences of Brain Lesions and Ablations as well as reports of specific experiments such as Witt's use of his observations of the changes in spiders' web construction as a means of identifying small quantities of hallucinogenic substances. Some of the papers seem entirely suitable for the purpose expressed in the preface, particularly those from popular science publications; for example, Selye's account of his work on the stress syndrome. Other papers, such as Guntrip's critique of Fairbairn's Theory of Schizoid Reactions, are for specialists alone. In general, while each paper is by itself worthy, it has gained nothing from being selected to appear in this book.

However, the book has value as a collection of some of the most important articles in the recent literature, taken from a wide range of psychological, psychosomatic and psychiatric journals. It is doubtful if it could appear in the undergraduate's book list except where his reading of it could be suitably monitored. It could best be placed in a hospital library where it can serve the postgraduate clinical psychologist and psychiatrist.

R. M. MOWBRAY

Psycho-Analysis and Contemporary Thought. By D. W. Winnicott, John Bowlby, Ilse Hellman, Marion Milner, Roger Money-Kyrle, Elliot Jaques and Joan Riviere, with an introduction by Sylvia Payne. Edited by J. D. Sutherland. London: Hogarth Press. 1958. Pp. 149. Price 25*s*.

This short book constitutes the published version of six public lectures given by members of the British Psycho-Analytical Society during the celebration of Freud's centenary in 1956. To these as introduction and tail piece have been added Dr Payne's introductory remarks as President of the Society and a note on Freud as a person by Joan Riviere.

The six main authors have dealt respectively with psychoanalysis in relation to the sense of guilt, child care, the teacher, art, philosophy and the then current economic crisis. As befits such lectures, and the resultant book, both intended for the general reader, the authors have confined themselves in the main to simple, often persuasive, exposition of familiar viewpoints. This reviewer found Dr Winnicot's subtle and sensitive paper and Mrs Milner's article the most worth re-reading and can readily understand that those who heard these lectures would want them to be in print. The other papers are more ephemeral.

The book should help in dispelling the stereotype of the psycho-analytic thinker as gloomy, ethically naïve, incurably dogmatic and inclined to use awful verbiage. It will not satisfy those who feel that, difficult as this is to do skilfully, any psychological position should be presented as potentially changeable in the light of well-described evidence. Would not even the general public have been refreshed and stimulated by more indication of the enormous changes that have in fact been made in psycho-analytic theory and practice since Breuer and Freud first saw the value of 'chimney sweeping'?

JOAN WYNN REEVES

Human Groups. By W. J. H. Sprott. Harmondsworth: Penguin. 1958. Pp. 219. 3*s*. 6*d*.

Contributors to the Pelican Psychology Series must somehow mediate between the professional worker and a large public, without displeasing either. Prof. Sprott qualifies as an excellent go-between. He has the necessary light touch, a clear 'no-nonsense' style and a respect for facts. Furthermore, *Human Groups*, with its 229 references, will be a welcome survey for those whose main preoccupation is in some other corner of psychology. The author himself is of somewhat mixed allegiance and this appears as a virtue when the work lies on the boundary of several properties, even if it does lead him to sit a little warily on the dividing fences.

The first three chapters define the field, those who work in it and the methods they use; the next three deal with permanent small groups, the family, village and neighbourhood; a further three with experimental groups and the final two with 'Crowds and Other Groups' and 'Applications' respectively.

Every book of this kind must explain how the individual and the group are conceptually distinct, but at the same time part of one social process. This is skilfully done, with emphasis on the development of self-awareness, the secondary need for the approval of others, the evaluation of our own responses with simultaneous prediction of the riposte, and the social origins of morality. The influence of G. H. Mead is perhaps the most noticeable, but the range is wide. It is a pity that this explanation is pieced together without any reduction to the more basic individual processes explored by experimental psychologists. Perhaps such an approach would estrange the general reader—but it is courting him a little unashamedly to refer to Homan's *The Human Group* in adulatory terms and to complain at the 'distressing jargon' of much other work.

There is a lack of balance between methods and findings. The latter are served up in quantity (though almost *too* liberally sprinkled with grains of salt for the taste of those who are ego-involved). For the former, the layman may wonder whether research involves anything more systematic than snooping, or having a really ingenious idea for tricking a group into thinking it is doing something that it is not. The method of *observation* gets a page and a half without any discussion of sampling, interviewing, participant observation or interaction recording. The *experiment* fares better, but mainly by way of example. The treatment is of the 'all things shall be kept constant' kind, and devices for achieving this that are important in psychology, indeed, in understanding the book—are not explained. Control groups are an example. These two methods are followed by three others—Sociometry, Field Theory and G. C. Homan's. The last two, certainly, are hardly methods. To imply that they are alternatives to observation and experiment is misleading, even if they are called 'approaches'.

More on methods, then; a separate section on theories (the author knows a lot about Parsons, whom he never mentions); and rather less about the so-far unco-ordinated artificial group material—would right the balance for this reader. But reviewers are socially conditioned to cavil, and these remarks should not decry the whole by any means. The book represents an important and difficult task well done.

TERENCE LEE

After-Effects in the Behaviour of Mice. By KR HOLT HANSEN. Copenhagen: Munksgaard. 1956. Pp. 435. 58*s.*

Aggression. By JOHN PAUL SCOTT. University of Chicago Press. London: Cambridge University Press. 1958. Pp. xi+150. 28*s.*

The book *After-effects* gives an account of original experiments with mice in new types of mazes called V-devices. Being a thesis, it states everything explicitly and is replete with definitions, diagrams, tables, summaries and individual records for hundreds of animals. Though the treatment is competent and, as far as novelty allows, clear, a reader well versed in the American tradition of animal psychology, but not the European, will not find it easy going. A striking contrast with American work is the fact that the author observes his animals keenly, and seems genuinely fond of them. This is a feature of which the reviewer heartily approves. In research on bar-pressing, as well as on maze learning, a lot of useful information is neglected if the animals are allowed to clock up 'responses' on the recording apparatus, unobserved by human eye. The data which the experimenter, and later the theorist, are thus given, are a pre-selected and sometimes unrealistic sample of behaviour. After generations of such research, problems and their treatment become artificial.

The author mentions four methods of treating an experiment: (A) non-interpretative registration [a factual record]; (B) immediate interpretative description [a frankly anthropomorphic account, made on the spot, of the animal's apparent emotions, thoughts, etc.]; (C) statistical analysis; and (D) causal interpretation [theories]. He claims that B is put forward merely as a working hypothesis and that it does not contaminate A and C. Perhaps unexpectedly, he is right—a strict behaviourist can ignore the subjective parts and consider only the facts.

In addition to the usual record of left or right turn, etc. (macrobehaviour), the author deals with microbehaviour in the V-field [the area around the choice point], e.g. the animal stops or almost goes left then retraces and goes left, turns its head, etc. He shows that the types of microbehaviour, far from occurring at random, are statistically related to the various forms of macrobehaviour, and are dependent on the task and on the animal's experience prior to reaching the choice point. Certain items of microbehaviour often go together while others are almost mutually exclusive.

The general outlook is gestalt, with emphasis on the animal's phenomenal world. In this connexion there are some touching drawings of mice illustrating 'disappointment, startle, surprise, nervous washing', etc. The types of microbehaviour are given tentative subjective interpretations, e.g. stopping at the V-field and turning the head in the direction not finally chosen gave the observer the impression that the animal was taking stock of the situation, looking around for something or other and speculating about something in a particular direction. Fanciful? The reviewer is by policy against this form of expression. But, in view of the frankness of the author, who is neither naïve nor ignorant of behaviourist work, and in view of the competence of the factual discussion, he feels that the temptation to dismiss such stuff off-hand should be resisted. While it is not likely that a positivist student will be converted to the European viewpoint there is much in this book to stimulate him, if he will only remove his blinkers for a while. Quite apart from the knotty question of animal consciousness, such a full account of behaviour is a refreshing change from the common stereotyped account of 'responses'.

Then there are reasoned attacks which a behaviourist should be prepared to counter. The learning curve is dismissed as only a rough measure, of little fundamental value. The law of effect is queried, since in many tasks the frequency of success rises from 50–75%, not to 100% as (according to the author) reinforcement theory demands. It is suggested that the behaviour is the result of a balance between two needs—hunger and the need 'to be up and about', and a theory of a Mendelian probability type is offered. The reviewer is not persuaded that the law of effect demands one dominant need, or that it cannot be made to cope with competing need if such are demonstrated. When the author himself accepts that 'need reduction' can promote the formation of an engram factor [strengthen an association] his objection to the law is seen to be only partial.

Aggression sets out to give a scientific account of aggressive behaviour, bringing the results of animal experiments, including those of the author on mice, to bear on human society. Various chapters deal in precis form with the psychological, physiological, genetic, social and ecological factors concerned with aggression—a wide field in which the author's powers are dissipated. No previous knowledge of these fields is required, but 'a common level of scientific competence' is assumed and the work is not intended as a popularization.

Though the text gives a superficial impression of rigour, the evidence offered frequently leads in a very tenuous manner to conclusions and recommendations on human behaviour. These are of a common-sense kind, the conclusions sometimes implausible and the recommendations sometimes impractical. For example, Pavlov found that a frequently repeated non-reinforced secondary stimulus tends to produce an inhibition. From this the author derives the principle of passive inhibition—learn not to fight by not fighting. It follows that grown-ups should not stimulate children to aggressive behaviour but give them peaceful surroundings. [How does one get other children to co-operate?] This treatment is 'relatively easy to do for a young baby'. [Such is not the reviewer's experience.] Again, if you nip a mouse's tail it will attack you, so pain is a stimulus for aggression. But 'no scientist wants to experiment on pain as a cause of human fighting by going out and punching the nose of the nearest bystander'. What scientist would suppose pain to be the important factor here?

The James–Lange theory of emotion is completely misrepresented. The notion of a threshold is sometimes used in a merely picturesque manner. No distinction is made between the threshold of stimulation for aggression and the intensity of aggressive behaviour resulting from stimulation. The persistent looseness of terminology can be illustrated: 'When a person "feels angry", this sensation comes from his hypothalamus and cerebral cortex, which have been stimulated to fight.'

The frustration-aggression hypothesis is rebutted by examples in which frustration does not lead to aggression. The fact that it is the aggressive behaviour in question which is being frustrated is not seen as making the examples inappropriate. Instinct as 'a driving force' is made into a straw man and demolished, while there is only one minor reference to Tinbergen.

The book contains a certain amount of interesting information, not-always-accurate snippets of Pavlov, Freud, James, Cannon, Mead, which cannot be found together in any other small volume. These may give the newcomer a bird's eye view of the subject but the lack of rigour prevents the book from being a reliable guide for the serious student.

These contrasting books illuminate two points which are by no means obvious. First, a wide knowledge, of the natural history type, and a perhaps sentimental interest in animals as creatures, not merely as bits of experimental apparatus, do not preclude a sound experimental technique. Second, a mere profession of behaviouristic rigour is no safeguard against the subtle intrusion of subjectiveness.

J. R. TROTTER

Traffic Accidents and Driver Characteristics. By SAULI HÄKKINEN. Helsinki: Finland's Institute of Technology, Scientific Researches No. 13. Pp. 199. 20*s*.

This is a doctoral thesis on accident proneness, translated into English and published as a whole. It is a good thesis, impeccable in method and representing a large amount of experimental work, although the nature of the topic studied has not lent itself to any great theoretical originality.

The foregoing sentences indicate the assets of the work. There is a modern review of the literature on the subject, considered from every point of view: there is a statistical examination of accident distributions from the Helsinki Transport Department: there is an account of a very large number of tests applied to the same employees appearing in the accident distributions, with an examination of the differences in scores found between high and low accident groups. There is also a factor analysis of the test scores including accident records. Finally the tests found most useful were applied to further groups in other kinds of work and generally similar results obtained. A particular virtue of the monograph is the international nature of the references and tests employed. The names of Russell Davis, Whitfield, McFarland, and Fleishman, for example, mingle in the bibliography in a thoroughly desirable fashion, and almost any reader is likely to find foreign references which will be new to him.

The results are not very novel, although probably timely since some scepticism about accident proneness has been expressed in the literature recently. The distributions were definitely different from Poisson ones, and high-accident individuals from one period had many accidents in the next even though differences in hazard were carefully eliminated from the analysis. Tests of intelligence or reaction time bore little relation to accidents: tests of co-ordination and of performance under stress did show significant results. But on analysing inter-correlations it appeared that tests predicting accidents did not necessarily correlate with each other, so that accident proneness is probably not a single entity but rather a label for a number of different departures from the ideal.

The thesis nature of the work does not make it rapid reading: every precaution taken is fully described, all the raw matrices of correlations given, and such details as the formula for calculating t are presented. Nor will the student of accidents find anything here to upset all his previous ideas. But all in all he should read this report.

DONALD E. BROADBENT

Statistics in Psychology and Education. By HENRY E. GARRETT. Fifth edition, 1958. London: Longmans. Pp. xii+478. 35*s*.

Substantial changes have been introduced in this new edition, the most immediately obvious of which is improved lay-out in a larger format. Part I, dealing with descriptive statistics, remains much as it was, but Part II has been largely rewritten to conform with present-day trends, particularly the increase in emphasis on probability and statistical inference. In this connexion it is worth noting that Dr Garrett has now almost entirely abandoned the use of the term 'reliability', preferring to write of the 'accuracy' of predictions, and of the 'significance' of statistical measures generally. Non-parametric methods are included for the first time: three of the simplest of these (in addition, of course, to chi-squared) are given brief treatment.

B. SEMEONOFF

Young Children in Hospital. By JAMES R. ROBERTSON. London: Tavistock Publications. 1958. Pp. xii+103. 4*s*. 6*d*.

This publication is a continuation of the work carried out by Dr Bowlby and Mr Robertson in their studies of the effects of maternal deprivation on children. The author, in this book, urges the need for reorganization of hospitals so that 'mothers-in' units will become an accepted part of hospitable life. The adverse effects of separation from the mother during hospitalization of the young child are discussed, and certain suggestions made to overcome these. A new film, 'Going to Hospital with Mother', which illustrates some of the recommendations contained in the text, is also available.

MARY COLLINS

Educational Research. Ed. by W. D. Wall. Vol. 1, No. 1. National Foundation for Educational Research in England and Wales. London: Newnes Educational Publishing Co. 5*s.* 6*d.* per issue: 16*s.* 6*d.* per annum (3 issues).

One always receives the first issue of the first volume of any periodical with interest. It is only over a period of years that one can judge the value of a periodical, but the first issue sets the course and offers a tempting sample of things to come. The purpose of *Educational Research* is to bridge 'the increasing gap between the worlds of research and the schools'. 'Contributions which make a readable synthesis of valid research on the practical problems posed by the English educational system' are invited. There is, the editor states, no intention of 'popularizing' or writing down.

The contents of this issue reflect editorial policy. Articles on the present concept of 'Intelligence' by Prof. P. E. Vernon, on Transfer at 11+, by A. Yates and D. A. Pidgeon, on the Wish to Learn by W. D. Wall, and on a comparison between English, Australian and Californian schools by D. A. Pidgeon, represent the 'comprehensive and authoritative statements of the principal facts and conclusions derived from research'. 'Certain topics of broad importance' are to be treated in serial fashion. An article on Teaching Children to Read, by J. M. Morris, is the first of such a series. Another feature is the annotated bibliography, the subject here being Arithmetic, the compiler J. B. Biggs. An abstract of a Swedish research completes the issue.

As the editor welcomes suggestions, we might make the following: (*a*) footnotes in larger type, (*b*) the inclusion of reviews of tests and books for the guidance of the class of readers the periodical is intended for, (*c*) printing the contents at the beginning of the issue as well as on the back cover.

Those who work in the field of educational research are very much aware of the 'gap' mentioned above. All such will welcome this new periodical, wish it success, and await future issues with interest.

JAMES MAXWELL

Teaching: a Psychological Analysis. By C. M. Fleming. London: Methuen. 1958. Pp. xiii+291. 28*s.*

Even a glance at this book shows it to have been aimed at the target of direct relevance to the job of teaching school children. Thus, in addition to the customary discussions of learning, motivation and the like, there is a chapter on 'the pupil in the classroom situation', another on 'craftsmanship in teaching' and there are sections on such topics as 'community resources' and 'administrative responsibilities', which latter even includes some practical advice on the purchase of text-books and tests.

The book, then, seems to be intended to serve as a very comprehensive guide to the beginner. However, it would be quite wrong to give the impression that it aspires only to offer useful hints and suggestions. There are two aims, over and above direct, practical relevance, which are just as important in the conception of the work. The first of these is to put every topic that is discussed into historical perspective; the second to emphasize the importance in all teaching of questions of human relations. The two are not left unrelated, for Dr Fleming repeatedly sets herself the task of showing how the present concern over social issues has grown out of earlier preoccupations.

There are few people who would question the value for the student of historical perspective; but this is quite a short book and a wide-ranging one. The historical surveys, then, are inevitably very brief, and there does arise doubt about their effectiveness for the average student in training as a teacher. Dr Fleming tells us at the start that she makes no apology 'for pausing...at various points to remind the reader of the contributions made by certain influential thinkers of the past'; and indeed no apology is due to any reader for whom Dr Fleming's surveys can fittingly be called 'reminders'. But, while some of the writings which she passes in rapid review may be, as she claims, 'so well-known that it is unnecessary to reproduce their content', this is certainly not true of all or most of them. More than this, however: one cannot expect the majority of those for whom the book seems to be intended to be familiar with the terminology of recent psychological inquiry; and Dr Fleming does not spare her readers in this respect. On the contrary, she offers a number of summarizing tables in which appear lists of the 'representative phrases' of the various writers: a formidable array. The supporting text is often so condensed that it is hard to see how the student can avoid bewilderment without extensive further reading; and in such a situation there arises always the danger of resort to rote learning.

Dr Fleming has perhaps not resisted quite sternly enough the temptation to try to achieve too much. But the book has a great deal to say, both by way of recording facts and urging values, that merits the attention of any teacher.

MARGARET DONALDSON

The People of Ship Street. By MADELINE KERR. London: Routledge and Kegan Paul. 1958. Pp. vii+215. 23*s.*

This account of a community in the slums of Liverpool is written simply. It breaks down the results of a study of sixty-one families under appropriate and sensible headings; children, adolescence, sex and marriage, clothes and food, crime, superstition and religion are dealt with chapter by chapter and make interesting reading. But this community is different; it is of Irish descent, it does not display the adaptive characteristics said to be necessary in an urban society; as an institution, it is immature in the sense that nobody ever seems to grow out of it.

The main reason for this is the power of the mother. Children do get married eventually, but they do not move away from the influence of their families, in particular, from their 'Mums'. If husbands do not choose their work in accord with the wishes of the Mum, they may find that their job has been changed for them. This tie to the mother seems to the author to be stronger than all other love relations; it generates a lack of independence; it leads to non-integration of the personality which results in piece-meal reactions and inability to plan. It results in a matrilocal society (at odds with the rest of our society, a patriarchal one), wherein there is lasting role inhibition, an excessive limitation of the number of roles which can be taken up, arresting personality development. Excess energy is absorbed in activities of crime and delinquency.

This attention to the social and societal determinants of individual immaturity seems most worthy and deserving of a much larger exposition than appears to be possible in this series of publications. As it is, the specialist will feel that in many places, the treatment is too slight. The clinician will tend to regard the evidence in the sections on projective testing as rather scanty. The same could be said for the analysis of the social structure. Here, in an effort to make the book more readable, none of the social statistics is treated in any kind of tabular form which, by its compactness alone, could have given the picture quite speedily to the interested eye.

The author is not entirely to blame for this. This kind of defect stems from the present organization of the behavioural sciences whereby a publication dealing with personal, social and cultural variables cannot be comprehensive without being vast. If any simple treatment is adopted, the study lacks depth and most of its value can be lost. It is unfortunate that from a research covering five years' intensive field study, it is the 'general' reader who stands to gain the most. GEORGE M. STALKER

Der Mann-Zeichen-Test. By HERMANN ZILER. Münster: Aschendorff. 1958. Pp. 51. DM. 6.75.

Ziler's 'Man-Drawing' test is very similar to Goodenough's. Minor differences are that Ziler gives fewer examples to guide the beginner in scoring, and that the scale does not extend sufficiently far down to cover the very immature child. On the other hand, Ziler's instructions are briefer than Goodenough's, and possibly more generally acceptable. The pamphlet is written in very simple language which should be readily understood by the average teacher. While there is considerable stress on statistical treatment, this too is kept on an elementary level. An aptly worded warning is given: '*Ein Test ist kein Test*', i.e. that the test should be applied only as part of a battery. Ziler finds a sex difference favouring girls for some details, and in general the older girls scored better than the boys, who tended to regard the test as somewhat trivial, while the girls got on with the task. A further interesting point is that in giving the test to fifteen hard-of-hearing children Ziler found that they either did not draw ears at all or drew them specially large, with strong lines. Goodenough does not note this, although she explicitly recommends the test for children with hearing difficulties. M. BEVERIDGE

Explorations in Social Psychiatry. By ALEXANDER H. LEIGHTON, JOHN A. CLAUSEN and ROBERT N. WILSON. London: Tavistock Publications. 1958. Pp. 452. 40*s.*

Each of the thirteen papers in this volume is an independent expedition into the little-known territory of social life where mental illness is propagated, nurtured, lived with, classified and treated. The title may be misleading to British readers, who would consider it a contribution to psychopathology, reserving the term psychiatry for the art of applying this basic knowledge. The editors would be the first to acknowledge that so little is known of the area that even the sadly devalued word 'integration' would be a bit premature, but the book is the better for being free of unwarranted theoretical posturings, and the overall impression is that the diverse research efforts of the authors, ranging from clinical insight to large-sample ecology, will all make their mark on the ultimate map.

Probably the best paper at the clinical insight end, is Kubie's exposition of his psychodynamic view of aetiology, and of what it implies for the problem of demarcating health from illness. The healthy person is one whose behaviour is 'flexible' because relatively small areas of his life are dominated by unconscious innaccessible processes. At the opposite pole, Eleanor Leacock analyses the methods and findings of ecological studies. Those who are content with the assumption that most disorders are spread thicker on the ground as they go from the green belt to the slums will find that a great deal more sophistication has been introduced into this area in recent years. A single criticism of an excellent article is that it neglects to mention the highly relevant 'about face' that has taken place with regard to desocialization and 'anomie' among students of the urban community. Gruenberg's chapter on Socially Shared Psychopathology is a fascinating review of the literature which subsumes dancing manias, the South Sea Bubble and the contagion of delusional systems between two friends or relations. Clausen compares the results of current researches on adolescent drug use. Hinkle and Wolff report a statistical (hardly experimental!) investigation of the incidence of illness in the life history. Much detail is given of the episodic distributions, but unfortunately the data from which they adduce their main conclusion is not presented fully. Illness of all kinds, they aver, clustered in time but spread over all body-systems of a single organism, is associated not only with infantile but also with adult *psychological* stress. Other topics are the effect of group atmosphere, in this case one described as *disparagement*, upon the mental health of a group; the varieties of response to bereavement; the concept of health in psychiatry and the problems of establishing a community psychiatric clinic.

There is no claim to have charted the whole territory. For example, Tyhurst describes paranoid patterns in detail, but there is no equivalent treatment of the other major groups. Homosexuality and alcoholism are not considered and deprivation is mentioned only briefly and with unwarranted dogmatism by Kubie.

A brave attempt at co-ordination is made by the editors, who intersperse a commentary before and after every chapter. But this is not altogether successful. In place of theoretical integration they often achieve only cross-reference and, alas, repetition. The effect is reminiscent of an American club sandwich, which one longs to reduce to edible proportions by dispensing with some of the bread.

Psychologists may be surprised that a book of this kind can be produced by an alliance of psychiatrists and sociologists with hardly a reference to themselves. This is only partly professional blinkerism—very few psychologists are in the field. Those wishing to enter could be safely advised at least to start from here. Plenty of virgin territory awaits them.

TERENCE LEE

Relaţiile dintre Gândire şi Limbaj in Ontogeneză. (The Relationships between Thinking and Language in Ontogenesis.) By T. S. CAZACU. Ed. Acad. R.P.R. 1957. Pp. 508.

Revista de Psihologie, 3. III. Inst de Psihol, Bucharest. 1958.

The central aim of Mrs Cazacu's study is to find out whether 'the child's thinking, and consequently his knowledge, are influenced—in a positive or a negative manner—by his gradual acquisition of language'. For this purpose she has devised a series of tests—based mainly on the verbal association technique—which she applied individually to 120 children from Bucharest ranging in their age from three to seven. Of a particular importance are the 'genus' tests designed to investigate into the child's abstract thinking. Interviewing and observation are used as additional methods.

The most important findings can be formulated as follows: (1) Even at the age of three children react 'meaningfully' to the word stimulus. (2) Their capacity for 'correct' association is closely related to their degree of familiarity with the words stimuli. (3) Pre-school children make in various degrees use of 'integrating criteria' in their thought processes, viz. they are successful in finding the word which designates the 'genus' of a group of objects. Special attention is paid to the variation of this trait according to age. The 'integrating criteria', based in the low age-groups on concrete perceptual factors, become, in the higher age-groups, more and more abstract and formal. The age of five is a crucial point in this development. The material collected throws light on other aspects in the child's thinking, such as word contamination, the formation of meaning, and the development of grammatical forms.

This study deserves to be praised on many accounts: the amount of work involved is impressive, and the author's experimental skill is remarkable. It is, however, open to criticism on one important point which refers to what the author calls the 'dialectical' relationship between language and thinking. This implies, first, that language influences the child's thought processes by directing and fixing them in a certain manner; and secondly, that thinking itself acts upon language as a selecting and organizing factor. But how was the child's thinking in its non-verbal form assessed? Admittedly, the author makes frequent references to such a phenomenon, particularly when she deals with the psychological roots of word contamination, but its existence is assumed rather than demonstrated. It seems that such a difficulty could have been at least partly avoided by a clear-cut definition of the basic terms 'thinking' and 'language', and by a stronger emphasis on the non-verbal aspect of the methods used. It is a pity, for instance, that the author did not describe more clearly her main experimental device, 'The Picture Cabinet', and particularly the bearing of this on the main thesis of the book. As it is, this study, however valuable, cannot go far beyond an investigation into the child's capacity for verbalizing his thought processes which, I feel, is not exactly what Mrs Cazacu intended.

Revista de Psihologie is the official Roumanian periodical in the field of psychology. As compared with the two pre-war periodicals this one shows a strong physiological orientation and a strong influence coming from Russian psychology.

Z. BARBU

Guide to using the Coloured Progressive Matrices, Sets A, Ab, B. By J. C. RAVEN. London: Lewis; Harrap. New edition, 1958. Pp. 40. 5*s*.

Guide to using the Mill Hill Vocabulary Scale with the Progressive Matrices scales. By J. C. RAVEN. London: Lewis. New edition, 1958. Pp. 64. 5*s*. 6*d*.

Earlier versions of these Guides have hitherto been supplied with specimen copies of the tests to which they relate. They have been subject to periodic revision, at about three-year intervals, and users aware of this fact should not assume that the present editions, now available as separate publications, are just another pair of revised reprints. Each of the present booklets has been entirely rewritten. They now provide a very full and carefully documented account of a testing approach which has been remarkably widely and successfully used, placing it in its historical context, with sections on construction and standardization, and much previously unpublished experimental and clinical material.

It is not, perhaps, well enough known that the Coloured Progressive Matrices may be usefully applied to adults—particularly old people—as well as to children. Both tests, whether applied separately or—better still—together, yield important qualitative as well as quantitative information; the evaluation of both types is fully discussed in the Guides, which between them present and interpret test data covering the whole life-span from 5 to 85 years.

B. SEMEONOFF

A Textbook of Psychology. By D. O. HEBB. Philadelphia and London: W. B. Saunders. 1958. Pp. x+276. 31*s*. 6*d*.

This is a fresh and rather exciting little book. About six years ago, Dr Hebb tells us in his Preface, he took a new look at the introductory course in psychology in his own University and found no text-book that covered what he wanted in subject-matter. Many teachers of psychology have had the same experience but few of us have responded with the fortitude of Dr Hebb. He wrote his own.

Psychology, to Dr Hebb, '...is fundamentally a biological science, not a social science, not a profession'. This means that it should take its departure from anatomy and physiology, zoology and genetics, and endeavour to provide a coherent account and explanation of the general mechanisms of adaptation. Now certain of these mechanisms involve what it has been customory to call mental activity, e.g. perceiving, imagining and thinking. Hence psychology must include within its scope processes of an essentially subjective character. But it is far more rewarding, Dr Hebb suggests, to consider the evolution and neurological basis of such processes than merely to examine and classify their formal characteristics. Hence he has little use for introspective analysis for its own sake. For him, modern psychology really began with Thorndike and Binet.

All this would seem to place Dr Hebb firmly in the tradition of American functional psychology—a tradition which grew out of the work of such men as James, Ladd, Dewey and Angell, and which shows gratifying signs of outlasting its monstrous offshoot, Behaviourism. But Dr Hebb is more of an experimentalist than were his 'functional' predecessors and it is this above all that makes his text-book something very much more than a neo-Darwinian tract. Although Dr Hebb's sense of evidence is sometimes shaky (most notably on issues of early learning), he is at bottom a man who believes, with Pavlov, that theory must be checked at every point against the actual findings of experiment. And, unlike Pavlov, he extends this provision to his own theories too.

It will certainly be raised against Dr Hebb that he is 'just another reductionist' aiming to translate the unique realities of feeling and experience into neurophysiological circuitry, probably bogus. This would be both unkind and unfair. Like every psychologist worth his salt, Dr Hebb has made it his business to find out what physiologists can tell us about the mechanisms of the brain and, unlike many, he really has found out a great deal. Moreover, he has applied this knowledge with originality and skill to psychological issues. But psychology, he insists, can never become a branch of physiology. 'We cannot escape the need of large-scale units of analysis, nor the need of special methods of behavioral study on which such analysis is based' (p. 264). Neurological constructs are no more solid, no more 'real', than psychological constructs and psychological methods may, indeed, come to play an important part in the study of brain function itself. Let us, then, hear no more of this particular red herring. Psychologists who fear the taint of 'reductionism' would do better to stop complaining and start to learn some neurology.

It is scarcely possible to indicate in a short review the scope and content of this admirable book. Much of it can be regarded as a further essay in the vein of *The Organisation of Behavior*, as a continuation of the author's endeavour to build some sort of psychology from the known facts of comparative behaviour study and the largely unknown facts of cerebral neurology. But there are also shrewd and useful chapters on heredity and maturation, statistical thinking and problems of measurement, and on psychology and the scientific method. Above all, Dr Hebb is not afraid of real people and real life, if at times he seems curiously naïve with regard to psychiatric issues. Taken all round, this short book is one of the very few that can be recommended to students of psychology virtually without qualification. O. L. ZANGWILL

Social Psychiatry in Action: a Therapeutic Community. By H. A. Wilmer. Springfield, Ill.: Thomas; Oxford: Blackwell. 1958. Pp. xxiii + 373. 66*s*.

For ten months Dr Wilmer was Officer-in-charge of the 34-bed admission ward of the Naval Hospital at Oakland, California. During this time 939 unselected male patients spent on average about ten days in his ward, and the nursing staff was also constantly changing. In this highly unlikely setting Dr Wilmer decided to set up a therapeutic community, and his book is a detailed description of it.

Daily group therapy sessions were held with both patients and members of staff, while at the same time he markedly reduced the use of barbiturate and ataractic medication and placed an absolute prohibition on physical restraints, and the use of the Isolation Room. These measures were aimed at changing a busy admission ward from an impersonal transit camp devoted to classification and disposal into a place where the patient could begin to feel safe, and would learn to achieve some measure of control over his fears and his feelings of aggression. By thus ensuring a favourable initial response to the hospital and by trying to modify the patient's attitude towards his illness, it was hoped that the admission unit would be able to make a more positive therapeutic contribution.

One may very well feel that this short experiment was successful, but whatever the merits of the therapeutic community, the book itself leaves much to be desired. Dr Wilmer was fortunate in having the services of a research Psychologist and a senior Social Worker for his ward, but their work is not reported in his book. On the other hand, less pertinent information is often repeated unnecessarily, clinical illustrations proliferate long after the author has made his point and the graphs, the photographs, and even the thirty-page bibliography seem curiously irrelevant to the main purpose of the book. As a direct result the book is too long, too costly, and is often in danger of sounding pretentious. A simpler, more direct presentation would have been more in keeping with Dr Wilmer's obvious honesty of purpose. J. DREWERY

The Concept of Motivation. By R. S. Peters. London: Routledge and Kegan Paul. 1958. Pp. 166. 14*s*.

'The aim of this monograph', says its author is, 'to contribute to the development of theories of motivation by subjecting some leading examples of contemporary theories to conceptual analysis'. He criticizes the traditional assumption that there can be an all-inclusive theory of human behaviour from whose basic postulates answers to all forms of the question 'why does Jones do *X*?' will eventually be deduced.

For, he points out, the answers are often logically different and even logically exclusive: but theories of motivation tend to generalize one kind of explanation in order to answer quite different forms of the question. Perhaps the only generalization of which he approves is Allport's apophthegm—'motives are habits in the making'. But, he says, 'to explain everything a man does in terms of "motives" is logically inappropriate because it lumps together acting purposefully according to rules, acting with a goal but according to no established rules, and cases where something happens to a man and it is odd to say that he *acts* at all.' He concludes that 'psychology...has remained earth-bound in mazes and Skinner boxes because the highly general theories which it was hoped would emerge are *logically* impossible'.

Mr Peters's method is a kind of logical dentistry. With a deceptively simple use of words that makes the probing seem painless he removes the teeth from Freud, Pavlov, Hull and several other system-builders: for example, 'one of the merits of Hull's use of the concept (of drive) was that it gave false answers to limited questions'.

The moral which I read into this book is that in studying a limited sector of human and animal behaviour you may find that description and prediction is greatly simplified by some new concept, such as conditioned reflex, drive, or unconscious wish; to apply the new concept outside this sector is always intellectually tempting, and quite often satisfying, but sooner or later there comes a point at which it becomes strained, useless and even misleading. Mr Peters's criticisms of the system-builders who have reached this stage is a practical demonstration of a kind of failure which, I suspect, is highly probable for purely logical reasons. There is no scientific reason to doubt that *in theory* a complete explanation of human behaviour can be given in purely physiological terms. It is logically possible that there is another set of terms in which an equally comprehensive and equally usable explanation could be given. But I know of no other field of scientific study in which this has ever been the case. It is therefore highly improbable that psychologists will ever provide a single set of terms to rival in comprehensiveness the theoretically possible physiological explanation, although for limited sectors thay have provided and will continue to provide different and, as it were, incommensurate explanatory systems which are preferable to the physiological because they are more easily grasped and used.

NIGEL WALKER

Measurement and Statistics. A Basic Text Emphasizing Behavioral Science Applications. By VIRGINIA L. SENDERS. New York: Oxford University Press. 1958. Pp. xvi + 594. $6.00.

This is yet another attempt to present the elements of statistics in a text suited to the needs of the mathematically unsophisticated. To this end Mrs Senders has done three things. First, she introduces her students to the concepts of statistics through a 'Diagnostic self-test' designed to give them an idea of how much of their 'high-school mathematics' they have retained. This is an excellent device, one which the reviewer has himself used to advantage, but the present test is rather spoiled by there being at least five questions to which the answers supplied are either wrong or quoted to a degree of precision not warranted by the form in which the data are presented. Secondly, since the book is explicitly addressed to the student ('This book has been written for you. I might almost say it has been written *to* you') the author has thought it appropriate to adopt a colloquial, even slangy, style, along with occasional exhortations to 'Relax!' or 'Be of good cheer'.

Thirdly—and it is on this point that one feels one must chiefly take issue—the order in which topics are introduced has been determined on logical rather than psychological grounds. Priority is given to operations admissible at all levels (or 'scales') of measurement, beginning with the weakest level, the nominal. The result is that many difficult mathematical concepts, including those of information theory, are introduced in advance of the simplest arithmetical procedures, such as the calculation of a mean. The author takes care to concede this point; nevertheless, it is tacitly assumed (pp. 78 ff.) not only that the concept of the mean (at this stage designated 'average') is familiar, but that the student will be able to understand—or accept—the fact that an 'average' may be a 'best guess'.

To the reviewer it would appear that the author has committed a cardinal pedagogical error. Instead of building on the basis of what the student knows and can handle with confidence, she has failed to recognize that what the non-mathematical student finds difficult is not so much the actual manipulation of figures (although this too has its terrors), but the adoption of mathematical ways of thinking and the assimilation of mathematical concepts. Thus, as early as p. 25 the student is invited to follow a mathematical demonstration based on summation of a constant from 1 to N. Furthermore, probably as an inevitable consequence of the chosen order of exposition, many of the illustrative examples in the early part of the book are based on situations that are trivial or psychologically unreal, or both. Thus an example showing that in certain circumstances uncertainty in card-guessing can be reduced by 0·2203 bits occurs on p. 91, whereas measures of central tendency are not formally handled until about 60 pages later. Correlation is introduced on p. 133, in relation to ordinal measurement, but consideration of the reliability of a correlation co-efficient (or of any other statistic) is postponed until very near the end of the book.

In the book's favour it can be said that it is well up to date in that many recently introduced non-parametric methods are included; that there is an adequate selection of examples for the student to work out for himself; and that—facetiousness apart—the writing is lively and refreshing. The author's enthusiasm for her subject is very evident, and one can imagine her communicating this enthusiasm to her students at Antioch College. Perhaps a student should have reviewed the book. B. SEMEONOFF

Eleven-year-olds Grow Up. By JOHN S. MACPHERSON. University of London Press. 1958. Pp. xxiii + 186. 15*s*.

This book is a publication of the Scottish Council for Research in Education. It is sponsored by the Mental Survey Committee but the collating of data and writing of the Report is due to their Research Officer, Mr MacPherson. The book is described in a subtitle as 'A seven-year follow-up of a representative sample from the 1947 Scottish Mental Survey'. The sample consists of 1208 school children selected from the original 70,805, all of whom were born in 1936. The method of selection, known as the Six-Day Sample, was based on the subjects' date of birth and the sample was found to be representative of the total group with respect to age, sex, test score, geographical area, size of family, parents' social and occupational status.

The chairman of the Committee writes, in his Preface, as follows. 'The report thus describes what happened to our "miniature Scotland" from 1947 to 1954. In particular it examines the relation between measured intelligence at age 11 and the future progress, educational and vocational, of the pupils in the sample. The educational progress of the pupils between age 11 and the age of leaving school is followed and the attempt is made to assess their happiness while at school. Light is thrown upon the problem of the premature leaver, and on the appeal of continuation classes to the young worker. The report further discusses occupational choice; change of occupation; relationships between boys' final choices and the occupational class of father; the effect of parental ambitions; and sex differences...leisure interests... attitudes to National Service....' This sounds a promising and well worthwhile project. Moreover, the author states that 'the closest parallel to this Follow-up is that of Prof. Terman in California, whose sample, however, comprised only gifted children with I.Q. over 135'. Terman's choice of title (*The Gifted Child Grows Up*) presumably inspired the present title. But there, in the opinion of the reviewer, the effective resemblance ends. *Eleven-year-olds Grow Up* is disappointingly prosaic and unstimulating. It contains innumerable tables and figures and lists in which x is compared with y and p with q—not forgetting the possible relationships of x with p and y with q.

The report represents a great deal of field work conscientiously carried out by school teachers, home visitors, civil servants and voluntary workers, and an immense amount of careful analysis and study of the results so obtained. Yet it does not finally add up to very much: little emerges in the way of interesting findings or fruitful hypotheses. The reader gains an impression of meticulous honesty and painstaking work, leading to a mammoth 'so what?'.

We are told, for instance, in 'Educational Aspects of the Follow-up' that 'there is a clear difference in average I.Q. between those who were successful [in their school courses] and those who were not. Thus for a group of pupils it may be possible to estimate how many would be successful in a particular course, but it would be quite impossible to predict the individual pupils.' In the chapter entitled 'High and Low Scorers' we read: 'The average size of family for the high scorers [over 131 I.Q.] was 2·3 and for the low scorers [under 78 I.Q.] 4·6. It is clear that the high scorer tends to come from the smaller families.' Again, in the chapter on 'Employment', when comparing Apprenticeship jobs with Unskilled Labour jobs, we learn that 'the average I.Q. of those who fall away from the Apprenticeship jobs is lower than that of those who continue in them, and that the average I.Q. of those who rise from the Unskilled group is higher than that of those who remain'.

The information purveyed, and copiously documented, throughout the book is exactly what is to be expected and what has many times been demonstrated. The reader feels that if all the results are in fact equally predictable and unexciting, then the book should be briefer. If, on the other hand, interesting conclusions are buried in the data, it is surely up to the author to unearth them. An illustration may clarify this latter point. MacPherson writes, when discussing 'Classification of Jobs': 'To some extent the argument appears to be a circular one. We test intelligence, then proceed to classify by ability levels the jobs entered. If we find that there is a relationship between job and intelligence test score, how should this be interpreted? Is it simply an indication of the validity of intelligence tests?....' The query is hackneyed and unhelpful. A constructive approach would be to ascertain the *proficiency* of the subjects (their I.Q. being already known) in the various jobs; and perhaps to examine the I.Q. *variance* within each occupational group. This, however, does not appear to be contemplated.

There are eighty-six Tables in the text; two Appendices; no Index. Appendix I consists of specimens of the Schedules used (Sociological, School, Home Visits and Employment Record), whilst Appendix II comprises a list of the M.A., C.A. and I.Q. for all members of the sample. The inclusion of the latter seems hard to justify: the reviewer would gladly exchange this list for an Index. The first Appendix, however, is intrinsically interesting and bears witness to the formidable amount of thought and attention to detail which has gone into the design of the Schedules, rendering them impressively clear and comprehensive.

A. W. HEIM

Productivity and Social Organisation: The Ahmedabad Experiment. By A. K. RICE. London: Tavistock. 1958. Pp. 298. 35*s*.

This book describes the impact of a technological innovation, and the organizational experiments and changes by which it was met, in a long-established Indian textile mill. One is at first struck by the very different cultural background industry has in India. Apparently only 1 or 2% of the population work as yet in modern mechanized industry, while three or even four different languages may be spoken in a single department of a mill. Industrial legislation has long aimed at reducing the employer's power over workers, so that discharge for anything but gross misconduct is so lengthy and difficult a process that it is almost never used. As for redundancy, the number of posts in a mill can be reduced only after written trade union agreement, or by a court award. Thus the immense unemployment is not allowed to become too directly the employer's weapon. Yet in fact these problems are basically familar to us; but being so much sharper in India they are much more likely to be taken explicitly into account in considering working relations.

Experiments and changes in three main areas are described. The first were in the automatic loom shed, where productivity was turning out to be no higher than with non-automatic looms. The new organization was based on a set of stated assumptions about work relations and satisfactions, drawn from studies of group behaviour by the Tavistock Institute, by Bion, and by others. Briefly it consisted of rebuilding the group of 29 workers, each with more or less specialized but very segmental jobs and all being concerned with the whole shed of 224 looms, into small groups of 6 or 7 workers, each group containing all the necessary skills and being responsible for all the work of a geographically compact group of perhaps sixty-four looms.

The second area of change was in non-automatic weaving, where the aim was to improve the quality of the work done. Changes here were on the same lines, but an additional difficulty was that there was no obvious 'new element'—as the automatic looms were—to help 'unfreeze' old ideas. As in the automatic shed, the soundness of the basic assumptions seemed to be confirmed by the response of the workers, and the new levels of output and quality.

The third area of change was in management structure. The other changes, though apparently involving only the shop floor, inevitably called for changes in the functions and responsibilities of members of management if they were to be maintained and exploited. Indeed, other studies of technological change have shown that it is more often at managerial than at worker level that the basic resistances are found. The arguments and analysis underlying the management changes at Ahmedabad form an important and thought-provoking section.

In effect the book is a case-history covering several years of considerable change in the life of a factory. It takes further the notions of Trist on 'socio-technical' systems, and underlines the fact—of current relevance—that increasing mechanization needs more rather than less attention to the social dimensions of production systems. All through there is a refreshingly adult and psychologically sophisticated approach to perennial problems of human organization. As a psychologist, one is left with the feeling that psychological research on organizational problems could be very much wider in approach and scope than many conceive it to be. One has to feel, too, how little of concrete knowledge psychology has so far contributed to this main area of life.

R. M. MCKENZIE

OTHER PUBLICATIONS RECEIVED

Freud and the Twentieth Century. Ed. by Benjamin Nelson. London: George Allen and Unwin. 1958. Pp. 314. 28*s*.

Tibetan Yoga and Secret Doctrines. Ed. by W. Y. Evans-Wentz. London: Oxford University Press. 1958. Pp. 389. 42*s*.

Colour Preference in Psychodiagnostics: a Survey. By Joh. de Zeeuw. The Hague: Staatsdrukkerij- en Uitgeverijbedrijf. 1957. Pp. 56. No price given.

Sampling Opinions: an Analysis of Survey Procedures. By F. F. Stephan and P. J. McCarthy. New York: Wiley. London: Chapman and Hall. 1958. Pp. 451. 96*s*.

The Integration of Educational Experiences. Ed. by N. B. Henry. University of Chicago Press. London: Cambridge University Press. 1958. Pp. 278. 30*s*.

Psychopathic Personalities. By Kurt Schneider. London: Cassell. 1958. Pp. 163. 18*s*.

Marriage and Personality. By A. Meares. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 157. 30*s*.

Psychology and Mental Health. Ed. by C. R. Strother. Washington: American Psychological Association. 1957. Pp. 154. No price given.

The Measurement and Appraisal of Adult Intelligence. Fourth ed. By D. Wechsler. London: Baillière, Tindall and Cox. 1958. Pp. 297. 40*s*.

The Physical Foundation of Biology. By W. M. Elsasser. London: Pergamon Press. 1958. Pp. 219. 30*s*.

Women. By Eustace Chesser. London: Jarrolds. 1958. Pp. 152. 10*s*. 6*d*.

Eduard C. Lindeman and Social Work Philosophy. By G. Konopka. Minnesota University Press. London: Oxford University Press. 1958. Pp. 220. 36*s*.

The New Chemotherapy in Mental Illness. Ed. by H. L. Gordon. New York: Philosophical Library. 1958. Pp. 762. $12.00.

Modern Clinical Psychiatry. Fifth ed. By A. P. Noyes and L. C. Kolb. London: W. B. Saunders. 1958. Pp. 694. 56*s*.

Personal Adjustment. By S. M. Jourard. New York and London: The Macmillan Company. 1958. Pp. 462. 38*s*. 6*d*.

Die Lebensphilosophie. By O. F. Bollnow. Berlin: Springer. 1958. Pp. 154. DM. 7.80.

Das Gewissen: Studien aus dem C. G. Jung-Institute Zürich. Bd. VII. By E. Blum, and others. Zürich and Stuttgart: Rascher. 1958. Pp. 208. No price given.

Personality: an Interdisciplinary Approach. By L. P. Thorpe and A. M. Schmuller. London: Van Nostrand. 1958. Pp. 368. 41*s*. 6*d*.

Cure de Sommeil Collective et Psychothérapie de Groupe. By H. Faure. Paris: Masson. 1956. Pp. 272. 2500 F. fr.

A New Pattern for Mental Health Services in a Children's Court. By H. B. Peck, M. Harrower and M. B. Peck. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 82. 25*s*.

Tribes Without Rulers: Studies in African Segmentary Systems. Ed. by J. Middleton and D. Tait. London: Routledge and Kegan Paul. 1958. Pp. 234. 28*s*.

The Idea of a Social Science: and its Relation to Philosophy. By P. Winch. London: Routledge and Kegan Paul. 1958. Pp. 143. 12*s*. 6*d*.

The Order of Presentation in Persuasion. Yale Studies in Attitude and Communication. Ed. by C. I. Hovland. Yale University Press. London: Oxford University Press. 1957. Pp. 192. 32*s*.

Introduction to Psychology. Second ed. By E. R. Hilgard. London: Methuen. 1958. Pp. 653. 55*s*.

Discussions on Child Development. Vol. 3. Ed. by J. M. Tanner and B. Inhelder. London: Tavistock. 1958. Pp. 223. 28*s*.

The Sources of Value. By S. C. Pepper. University of California Press. London: Cambridge University Press. 1958. Pp. 732. 64*s*.

Figural After-effects. By P. McEwan. Brit. J. Psychol. Monogr. Suppl. London: Cambridge University Press. 1958. Pp. 106. 22*s*. 6*d*.

Social Class and Mental Illness: a Community Study. By A. B. Hollingshead and F. C. Redlich. New York: Wiley. London: Chapman and Hall. 1958. Pp. 442. 60*s*.

City Life in Japan: Life in a Tokyo Ward. By R. P. Dore. London: Routledge and Kegan Paul. 1958. Pp. 472. 45*s*.

The Objective Psyche. By M. Fordham. London: Routledge and Kegan Paul. 1958. Pp. 214. 25*s*.

Jung and the Problem of Evil. By H. L. Philp. London: Rockliff. 1958. Pp. 271. 30*s.*

Jung and St Paul. By D. Cox. London: Longmans, Green and Co. 1959. Pp. 357. 32*s.*

Behavior and Psychological Man. By E. C. Tolman. University of California Press. London: Cambridge University Press. 1958. Pp. 269. 17*s.* 6*d.* (A paper-bound edition of *Collected Papers in Psychology*, 1951.)

World Health. By F. Brockington. Harmondsworth: Penguin. 1958. Pp. 405. 5*s.*

Patterns of Discovery: an Inquiry into the Conceptual Foundations of Science. By N. R. Hanson. London: Cambridge University Press. Pp. 241. 30*s.*

Extrasensory Perception and Personality Patterns. By G. R. Schmeidler and R. A. McConnell. Yale University Press. London: Oxford University Press. 1958. Pp. 136. 25*s.*

The Education of the Individual. By Alfred Adler. New York: Philosophical Library. 1958. Pp. 143. $3.50.

What, Then, is Man?: a Symposium of Theology, Psychology and Psychiatry. By various authors. Saint Louis, Missouri: Concordia. 1958. Pp. 354. $3.50.

Aviation Psychology in Western Europe. Ed. by J. B. Parry and S. D. Fokkema. Amsterdam: Swets and Zeitlinger. 1958. Pp. 64. 15*s.*

Greek Culture and the Ego. By A. Stokes. London: Tavistock. 1958. Pp. 101. 15*s.*

Tales from the Vienna Woods. By L. Koenig. London: Methuen. 1958. Pp. 159. 18*s.*

The Negro Personality. By B. P. Karon. New York: Springer. 1958. Pp. 184. $4.50.

Thrills and Regressions. By M. Balint. London: Hogarth. 1959. Pp. 148. 18*s.*

Flying Saucers: a Modern Myth of Things Seen in the Skies. By C. G. Jung. London: Routledge and Kegan Paul. 1959. Pp. 184. 14*s.*

Fundamentals of Measurement: Techniques and Practices. Oxford University Press. 1958. Pp. 413. 48*s.*

The Teaching and Learning of Psychotherapy. By R. Ekstein and R. Wallerstein. London: Imago. 1959. Pp. 334. 42*s.*

Live at Peace with Your Nerves. By W. C. Alvarez. Kingswood, Surrey: The World's Work Ltd. 1959. Pp. 267. 21*s.*

Psychiatry in General Practice. By J. A. Weijel. London: Van Nostrand. 1959. Pp. 208. 37*s.* 6*d.*

Behavior of Industrial Work Groups. By L. R. Sayles. New York: Wiley. London: Chapman and Hall. 1958. Pp. 182. 38*s.*

Ego Psychology and the Problem of Adaptation. By H. Hartmann. Trs. by D. Rapaport. London: Imago. 1959. Pp. 121. 21*s.* (This is the first number of the Journal of the American Psychoanalytic Association Monograph Series.)

Theoretical Criminology. By G. B. Vold. Oxford University Press. 1958. Pp. 334. 40*s.*

The Society of Captives: a Study of a Maximum Security Prison. By G. M. Sykes. Princeton University Press. London: Oxford University Press. 1958. Pp. 144. 30*s.*

Zum Problem der Sensumotorik. By Kurt Kohl. Frankfurt am Main: Kramer. 1956. Pp. 110. DM 6.00.

Mengenvergleich und Übung. By Egon Becker. Frankfurt am Main: Kramer. 1957. Pp. 63. DM 5.00

Über die Rolle der Bezugsbereiche bei der Lokalisation. By Kurt Müller. Frankfurt am Main: Kramer. 1957. Pp. 62. DM 5.00.

Instruktionsverständnis und Testleistung. By Manfred Sader. Frankfurt am Main: Kramer. 1957. Pp. 111. DM 6.00.

The Problem of Perception. By R. J. Hirst. London: Allen and Unwin. 1959. Pp. 330. 30*s.*

The Social Purposes of Education. By K. G. Collier. London: Routledge and Kegan Paul. 1959. Pp. 236. 21*s.*

The Art of Listening. By D. A. Barbara. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 201. 42*s.*

The Study of Man: Lindsay Memorial Lectures. By M. Polanyi. London: Routledge and Kegan Paul. 1959. Pp. 102. 7*s.* 6*d.*

Perspektiven der Persönlichkeits-theorie. Ed. by H. von Bracken and H. P. David. Bern and Stuttgart: Hans Huber. 1959. Pp. 319. DM 32.50. (The English edition of this book was reviewed in this *Journal*, vol. 48, pp. 310–11.)

Nonlinear Problems in Random Theory. By Norbert Wiener. New York: Wiley. London: Chapman and Hall. 1958. Pp. 131. 36*s.*

The Nature of the Self. By F. J. Mott. London: Allan Wingate. 1959. Pp. 316. 42*s.*

Early Education of the Mentally Retarded. By S. A. Kirk. University of Illinois Press. 1958. Pp. 216. $6.00.

Language and Psychology. By Samuel Reiss. New York: Philosophical Library. 1959. Pp. 299. $3.75.

VOLUME 50 AUGUST 1959 PART 3

THE SOLVING OF FIVE-LETTER ANAGRAM PROBLEMS

BY IAN M. L. HUNTER
Psychology Department, Edinburgh University

An experiment is reported involving 72 student subjects, a population of 144 different five-letter anagrams and a total of 864 solution times for analysis. Various characteristics of anagram solving are considered leading to a conception of the process as a recurring trial and check activity in which the nature and sequence of the tries are determined jointly by the thinker's set, linguistic knowledge, and preference for certain forms of letter position rearrangement. Regarding this last factor of letter location, solution is attained sooner if the thinker must change the location of a single letter (*a*) rather than the locations of several letters, (*b*) to a nearby rather than to a distant location, and (*c*) from an end to a middle rather than from a middle to an end location: ease of moving a single letter does not depend on the direction of the move. A subsidiary experiment shows that the relative difficulty of complex, as contrasted with simple, moves is less for students than for 15-year-olds than for 12-year-olds. This finding is interpreted in terms of interaction between letter location factors and varying levels of linguistic knowledge. It is suggested that improvement in overall anagram solving performance can be brought about only with learning which extends over a period of months or years.

I. INTRODUCTION

The purpose of this paper is twofold. First, to describe, in *formal* terms, the effect on anagram solution time of the way in which, irrespective of the particular letters involved, the letter order of the anagram is disarranged relative to the letter order of the solution word. Secondly, to present a generalized *functional* description of the processes involved in attaining anagram solutions. In biographical fact, this latter description derived from the results of the present experiment and of a pilot experiment conducted a year earlier; but it is convenient, for expository purposes, to begin with this functional description.

What does a person do while attempting to achieve the solution of an anagram? It would appear that his performance can be characterized as a sequence of more or less distinct activity phases. Each phase involves producing a potential solution or solution-component and assessing the appropriateness of this to the requirements of the problem. Thus, the speed with which he attains solution would depend on the nature of his proposals, their sequence, and the rapidity with which they are produced, one after the other, until one occurs which is judged appropriate to the problem as understood. These qualitative, sequential, and temporal characteristics of his proposals would, in their turn, depend on a number of interacting factors among which the following three would seem to be of particular importance.

(1) *Task-set*

To begin with, and throughout, the thinker must appreciate the general nature of his problem as being of the anagram class and, thereby, dispose himself to act in limited ways, even though his disposing is reflected but slightly or not at all in his conscious awareness: for example, he must not, as he may do under other circumstances, attempt to memorize the anagrams or treat them as acrostic forms. With literate adults, this general task-set is established so readily by brief verbal instructions that its presence and importance are easily overlooked. Its role is more evident when it is more difficult to establish, as when the thinker is a child or an illiterate; or in a cross-word puzzle situation where even a literate adult must discover that some 'clue' implies an anagram problem. The role of task-set is also more evident when its specificity is varied, as when the thinker is or is not told the category of the solution word beforehand (Hunter, 1956, 1960), or when a cross-word puzzle is presented by a constructor whose 'style' is or is not familiar to the thinker.

The role of task-set cannot be overemphasized since it pre-selects the range of the thinker's activity. The writer often asks students to solve the following problem. 'A man climbs a mountain path to the summit, turns, and descends the path. He goes up at an average speed of six miles per hour and comes down at an average speed of two miles per hour. Assuming he spends no time at the summit, what is his average overall travelling speed?' When presented in a casual manner which arouses no suspicions of a 'catch', this problem usually elicits the answer of four miles per hour. The writer then gives the correct answer and shows its derivation. 'But' the students are asked 'why the wrong answer?' They report having followed the habitual averaging procedure of halving the sum of the two numbers without having recognized that this procedure is inappropriate when the numbers are themselves averages. The point of this demonstration is that, in retrospecting thus, the students approximate to a formulation of the two-step process mentioned by Bruner (1958) as fundamental to much of thinking. 'The one [step] is a process of rule or structure learning, the second is a process whereby we recognize events as conforming to the structures we have learned' (p. 163). The students' error stems from their having miscategorized the problem and, as a result, transformed its data according to the wrong rules. People typically react to a strange scene, picture, person, story, or event by an 'effort after meaning', that is, by an attempt to relate it to their past experience: and once they have related the new and unfamiliar to the old and familiar, they can react to the novel situation by 'following the rules'. So, too, in facing a fresh problem. During their lifetime, people have met and solved many kinds of problem and, in doing so, have learned something of the procedures appropriate to their solution: if people can identify a new problem as being of this or that familiar kind, they are able to start dealing with it. Everyday observations provide many illustrations of the importance of recognizing the class to which a novel problem belongs. There are the precedents used by the lawyer and the administrator, the grounds on which the schoolchild either adds or multiplies the numbers in his arithmetic exercise, the confusion of the student who can carry out several statistical calculations but cannot decide which to apply to his present data. Laboratory illustrations of this factor are provided by some unpublished experiments of the writer. Hunter's (1957*a*) group experiment with students shows a strong 'atmosphere effect' in the attempted solving of a difficult problem concerned with proportional relations. More recently, he has attempted to elucidate the sources of this effect by observing the behaviour of individual students in tackling one or other version of this problem. Many instances of 'atmosphere' are found to derive from miscategorization of the problem: several forms of such erroneous categorizing are revealed but perhaps the commonest involves ignoring the constancy restrictions of the problem and treating it as demanding the construction of a three-term series. Again, the correct answer is most often given by those students who categorize the problem as one involving the familiar physical inter-relations which hold among the temperature, pressure, and volume of gasses. And yet again, some persistent students spend most of their solving time in trying out what amounts to one categorization of the problem after another, e.g., one student begins by saying 'let A, B and C be the squares on the sides of a right-angled triangle', finds this unhelpful, and carries out further 'substitutions' until, in his sixth, he arrives at a geometrical construction into which values of A, B and C can properly be fitted so as to generate the correct solution. As with 'atmosphere effects' shown by the students in the above problem situation, so with 'atmosphere effects' shown by children in taking three-term series problems (Hunter, 1957*b*). It has been suggested (Hunter, 1958) that these effects derive from the assimilation of the problem to the wrong class, from the application of the wrong rules, or, to be more specific, from treating serial relations either as superlative relations or as relations of inclusiveness.

In anagram solving, the importance of task-set is that it predisposes the thinker to perform certain broadly restricted activities and maintains these throughout the trial-and-check sequence. The most important of these pre-selected activities concern life-long verbal habits which must now be considered.

(2) *Linguistic knowledge*

The thinker's vocabulary is of obvious importance in anagram solving if for no other reason than that he must recognize the solution word as a word if he is to attain it as his solution. Differences in his familiarity with various words are also of importance: Mayzner & Tresselt (1958, 1959) confirm everyday observation by demonstrating that solution words of frequent occurrence in English (e.g. sugar) are, on the whole, attained sooner than words of lower frequency count (e.g. triad). In part, this finding may be due to the more common words being easier to recognize as appropriate once they have been produced—a circumstance found by cross-word puzzlers who sometimes attain a certain letter sequence but, because of limited vocabulary, do not recognize this sequence as being a word. In part, this finding may be due to a characteristic of recalling performance. At some points in the thinking sequence, some thinkers recall a succession of words which somewhat fulfill the requirements of a provisionally constructed skeleton of the sought-for solution; and it is known that, in the continuous controlled recall of words, the commonest words tend to be recalled earliest (Bousfield, Cohen & Silva, 1956). Lastly, the earlier attainment of more common words may, in part, be due to their containing those letter sequences which are of more common occurrence. Mayzner & Tresselt (1959) show that an important factor in a person's anagram solving is his knowledge of the transition probabilities between individual letters, that is, his knowledge of the frequencies with which this letter goes along with that in making up words. For example, the digram 'th' has a high frequency of occurrence and has been estimated by Pratt (1942) to occur some 168 times in 1000 words of 'normal text': the digram 'tf' has a low frequency of occurrence and appears only once in 1000 words of text. From the results of the Mayzner and Tresselt experiment, the present writer would draw the generalization that the combining of anagram letters into pairs of high digram frequency occurs earlier, on the whole, than the combining of letters into pairs of low digram frequency. Thus, it would seem that, if anagram solving is regarded as the successive permuting and combining of separate letters, it must be conceded that these combinations are carried out under the guidance of knowledge about the transition probabilities between the letters: the combinations (into groups of two, three, or more letters) which are, on the basis of past linguistic experience, most likely to occur are those which are produced earliest in the trial-and-check sequence while the production of less frequently encountered combinations are produced later. Indeed, the grouping of letters into sequences of very common occurrence may be carried out so promptly and the grouping adhered to so tenaciously that it might, for some purposes of analysis, be necessary to regard the thinker as working with units which are letter groups rather than individual letters.

At this point it is worth recalling a suggestion made by Hollingworth in 1935. As the person's verbal repertoire increases and as he comes to perceive more and more words as familiar units, then 'we might expect that words could be "built" from letters more readily or rapidly if these letters were presented mixed up in such a way that they did not already constitute a "word". Thus it ought to be easier to construct the word "intellectual" from the letters "alinletluetc" than from the letters "all in lettuce"' (pp. 299–300). Attempts to verify this expectation have led to ambiguous results and Mayzner & Tresselt (1959), who cite these attempts, propose that the conflicting findings might be resolved by rephrasing the expectation as follows. 'Well patterned letter groups are, in fact, more difficult to reorganize than loosely patterned groups, if patterning is defined in terms of transition probabilities' (p. 121). In brief, what is important may well be that the anagram's letters form (or do not form) commonly experienced letter sequences which must be broken up rather than the almost incidental fact that the anagram forms (or does not form) a word or phrase which must be broken up. With regard to this conclusion, it may be mentioned that exploratory experiments at Edinburgh indicate that speed of anagram solving is changed in a complex way when people have previously memorized the anagrams as nonsense words and presumably altered thereby the transition probabilities involved.

Closely related to the person's knowledge of transition probabilities between letters is his knowledge of the frequencies with which words have this or that vowel-consonant pattern. For example, thinkers given a five-letter anagram consisting of four consonants and one vowel almost always and immediately assume that the vowel will come in the centre of the solution word. Observational evidence alone seems to warrant the conclusion that the successive rearrangement of letters is carried out under the guidance of the thinker's statistical knowledge of both the structure of words and letter sequences. (It should be emphasized that, in the present context, 'knowledge' does not necessarily refer to relationships which can be verbalized by the thinker. For example, the writer 'knows' the positions of the keys on his typewriter in the sense that he can touch-type but not in the sense that he could draw an appropriately labelled diagram of the keyboard from memory.)

(3) *Letter locations in the anagram*

It has just been argued that, when a person is presented with an anagram, his task-set and his linguistic knowledge combine to yield a succession of letter groupings in which certain groupings are produced earlier than others. However, it is evident that some effect must be exercised by the locations of the letters in the anagram array for, if these locations played no part, then all anagrams sharing the same solution would, on the average, be solved in the same time—and such is not the case as will be demonstrated in the experiment reported below. Consider the extreme situation in which the thinker has, on the grounds of his linguistic knowledge, no preference for any particular letter grouping. An approximation to this situation may arise with linguistically inexperienced thinkers such as young children; or with linguistically experienced adults when, for example, they are faced with a choice between two digrams of equal frequency count. In such a situation of no preference, it is the locations of the letters in the anagram which seem likely to determine the letter grouping made since the thinker may well prefer some patterns of location moves to others. For example, he may try interchanging the locations of two adjacent letters earlier than he would try moving a letter to a non-adjacent location—in which case, the solution for which the first move is appropriate will be attained sooner than the solution for which the second move is appropriate. With people who are capable of solving anagrams, it is in practice unlikely that an anagram could be presented which would arouse no preference at all for one letter grouping rather than another: nevertheless, such preferences do seem to be influenced by the factor of letter location. It seems, for example, fairly evident that preference is given to letter groups already present in the anagram rather than to groups which must be constructed by combining letters that do not occupy adjacent locations. A clear instance of this occurs with an anagram which so resembles a word that the displaced letters are redundant (as are elusive printing errors) and do not impede immediate recognition of the word, e.g. the letter arrangement 'psycohlogy'. Such immediate word recognition may also delay solution as when the anagrams 'irato' and 'unclh' lead to difficulties because of their similarities to the words 'irate' and 'uncle' respectively. In general terms, it may be suggested that where an anagram contains a particular letter grouping, success will come sooner when this grouping is required for the solution word than when it is not because, in the latter case, its presence may set the thinker to try out letter combinations or to recall words with an inappropriate letter group as focus. A hypothetical case which illustrates this generalization is provided by considering the solution word 'ratio'. The following digrams have (see Pratt, 1942) much the same frequencies of linguistic occurrence: 'io' (31 times in 1000 words of text); 'ro' (31 times); 'ta' (29 times); and 'tr' (29 times). According to the above generalization and other things being equal (this is, let it be remembered, a hypothetical case), a thinker would be at an advantage in dealing with an anagram containing the first digram as compared with an anagram containing any of the other three.

It is the primary purpose of the experiment reported below to investigate speed of anagram solving as a function of the locations of the letters in the anagram, irrespective of the nature of the particular letters concerned.

II. MATERIALS AND PROCEDURE

The twelve anagram letter arrangements

Five-letter anagrams are used throughout and two classes of letter arrangement are contrasted. The *low scrambled* anagram contains three, and no more than three, letters in the same forward consecutive sequence as in the solution, and the solution can be derived by changing the location of only a single letter. Representing the solution as 1–2–3–4–5, the six anagrams conforming to these criteria are as follows. i, 1–2–3–5–4; ii, 2–1–3–4–5; iii, 1–3–4–5–2; iv, 4–1–2–3–5; v, 1–5–2–3–4; vi, 2–3–4–1–5. The *high scrambled* anagram maximally disarranges the letters of the solution word. The six following arrangements are used (note that one arrangement occurs twice). I, 3–1–5–2–4; II, 3–5–1–4–2; III*a*, 5–2–4–1–3; III*b*, 5–2–4–1–3; IV, 5–3–2–4–1; V, 4–2–5–3–1. Taking these twelve arrangements as a whole, the solution positions 1, 2, 3, 4 and 5 are distributed fairly evenly throughout each letter position of the anagrams. Thus a person taking all twelve anagrams could not develop any helpful set for one or a few forms of letter rearrangement.

The twelve solution words

These are: BRICK, WATER, ABOUT, SCARF, UNCLE, RATIO, LUNCH, OTHER, HOUSE, DEPTH, CHAIR and IMAGE. Note that each word occurs fairly commonly in the English language, contains no letter more than once, and permits of one solution only. Together, the words represent a variety of meaning categories and vowel-consonant patterns, that is, some contain three vowels, some two, and some only one, and the locations of these vowels are as varied as possible. Thus within the limits of the experiment and except for the critical variable (amount and kind of scrambling) an attempt is made to avoid sampling error by having both the solution words and the letter arrangements of the anagrams as varied as possible.

The twelve anagram lists

Each list contains twelve anagrams and is to be taken by a different person. In each, the twelve solution words always occur in exactly the order shown above: BRICK always comes first, WATER always second, ABOUT always third, and so on. Likewise, the anagram letter arrangements always occur in the following order: ii, III*a*, i, II, iv, iii, IV, I, vi, V, v and III*b*. But this order, although constant, is varied in cyclic fashion so that one list starts with BRICK in arrangement ii, another with BRICK in arrangement III*a*, another with BRICK in arrangement i, another with BRICK in arrangement II, and so on. Over the twelve lists as a whole, each anagram arrangement occurs once in each serial position so that every solution word undergoes all twelve forms of disarrangement. Such counterbalancing enables conclusions to be drawn about the average effect of each letter arrangement on speed of solving—independently of any effect due to a particular solution word or to practice. (However, conclusions regarding solution time as a function of particular solution words are confounded in so far as practice effects are present.) Each list is preceded by four practice anagrams, i.e. RLGI, EOGRAN, AEBOV and EEPHS. These are identical for all twelve lists and enable both subjects and experimenters to become acquainted with their respective tasks. Each list is typed on a foolscap sheet of paper. The anagrams, one below another, are in capital letters with a single space between letters and four spaces between anagrams. Each list is numbered from 1 to 12, and every anagram has a letter or number (not visible to the subject) identifying the particular letter arrangement involved.

Procedure

The experiment was conducted with seven class sections, containing about 20 students each, as part of the laboratory work of First Ordinary Psychology. Students worked in pairs, one acting as experimenter (E), the other as subject (S). Each pair was provided with an anagram list, a stop-watch and a movable aperture, i.e. a strip of hardboard large enough to cover all the anagrams except one which appeared through a slit in the centre of the board. All students had had previous experience of using both stop-watches and movable apertures, and all S's were selected from native Britishers. Students were not told the purpose of the experiment beforehand. They were reminded of what an anagram was and told that each S would be presented with 16 of them, one at a time, and would give the solution to each as quickly as possible. Solutions would all be familiar words in the sense that no student would be tempted to consult a dictionary for confirmation that the solution was, indeed, a word; furthermore, each solution would be one single word, and words containing two distinct parts, such as 'Z-bend', would not be permitted. (In pilot studies, 'throe' had occasionally been given instead of 'other' and 'U-boat' instead of 'about'. The above instructions were intended to preclude these solutions: and, in fact, these solutions were not proposed by any S in this experiment.) The E was told to alert his S, start the stop-watch simultaneously with exposing the anagram, observe the movement of the watch's hand, and note the time (to the nearest second) at which the solution was given: he was told to keep the watch running for a short time after a solution had been given in case this solution was recognized by S as being wrong and followed up by the correct solution. If no solution were given in a 30 sec. exposure, E was to record a 'fail'. The time elapsing between exposures of successive anagrams was to be left flexible so that the experiment could be conducted at a pace which flustered neither S nor E. After the experiment was completed, the students were invited to offer introspective comments on their anagram solving experiences.

III. RESULTS CONCERNING LETTER LOCATIONS

Results are available from 72 S's. Each of the twelve lists is taken by six S's, three of whom take the lists in the forward order (i.e. from top to bottom) and three in the reverse order. Excluding the practice anagrams, there is, therefore, a total of 864 solving attempts for analysis.

Low versus high scrambled anagrams

For each S, find the six anagrams solved in the shortest times (allowing for ties where necessary), and how many of these six are low scrambled. If low scrambled anagrams are more rapidly solved than high, this number should exceed the chance value of 3. In fact, this number is 2 (for one S), 3 (for 17 S's), 3·5 (for 3 S's), 4 (24 S's), 4·5 (2 S's), 5 (23 S's) and 6 (2 S's). The average value for all 72 S's is 4·1 out of 6. For purposes of statistical assessment, the binomial test (Siegel, 1956, pp. 36–42) is applied to the following data: 54 S's score more than 3 while only one S scores less than 3. The result is very highly significant and it can be concluded that low scrambled anagrams require, on the average, a shorter solving time than high scrambled ones.

Comparisons of the different anagram arrangements

The overall solving times required for different anagram arrangements are compared by considering two arrangements at a time. For example, in comparing iii and v, the total number of S's is found who, in their individual records, solve iii in less time than v: the total number is also found of S's who solve v in less time than iii: those S's are ignored who take the same time to solve both iii and v and, of necessity, who fail to solve both in the 30 sec. allowed. The two totals obtained should be approximately equal if, to take the null hypothesis, there is no difference in solving time due to the difference between the two anagram arrangements. So the binomial test is applied to discover whether these two totals differ from equality to a greater than chance extent. A two-tailed test of significance is used throughout since no predictions are being tested. In fact, 43 S's solve iii in less time than they solve v, while 21 S's solve v in less time than iii. Since these totals differ significantly from equality ($P = 0\cdot0086$), the conclusion is that, over all, iii is solved in less time than v. To furnish an empirical check on this statistical procedure, arrangement III is presented twice to each S. The number of S's solving III*a* in less time than III*b* is 26, while 27 solve III*b* in less time than III*a*. The difference is, as might be expected, negligible.

When the different anagram arrangements are compared as indicated above, differences between solving times emerge as follows.

The five high scrambled arrangements. Arrangement IV requires a significantly shorter time than any of the others. Also, I, II, III and V do not differ significantly among themselves. The P value for the difference between IV and III*b* is 0·0080. The lowest P value for the difference between III*b* and any of the remaining arrangements is 0·6030 between III*b* and V.

High versus low scrambled arrangements. Except for IV, all high scrambled arrangements require significantly longer solving times than any of the low scrambled arrangements. This confirms the conclusion drawn earlier. However, there is no significant difference between IV and either v or vi, each of these three arrangements being of almost identical difficulty.

The six low scrambled arrangements. The approximate order from easiest to most difficult is i, ii, iii, iv, v and vi. Within this order, none of the adjacent arrangements differs significantly from each other: the P value for the difference between i and ii is 0·4122, between ii and iii is 0·7840, between iii and iv is unity, between iv and v is 0·2006, and between v and vi is unity. Significant differences do, however, exist between non-adjacent arrangements as follows. Between i and iii ($P = 0\cdot0001$) and, *a fortiori*, between i and iv, v and vi. Between ii and iv ($P = 0\cdot0332$) and, *a fortiori*, between ii and both v and vi. Between iii and v ($P = 0\cdot0086$) and, *a fortiori*, between iii and vi. Lastly, between iv and vi ($P = 0\cdot0086$).

With the exception of arrangement IV (which will shortly be discussed), the above results suggest the following conclusions. Note that these conclusions are of a formal kind in that they merely state overall relations between solving time and the locations of the letters in the anagram: they do not, as did the Introduction, attempt to specify the psychological processes which bring these relations about. (1) A solution requiring only one position move is attained earlier than a solution requiring several position moves. (2) When several position moves are required, there are no differences in the times taken to attain different kinds of multiple moves. (3) Where only one position move is required, there are differences in the speed with which different kinds of single move are proposed. This third conclusion may be elaborated as follows. (3*a*) A shorter solving time is not entailed if the displaced letter must be moved backward rather than forward. This follows from the non-significant differences between i and ii (1–2–3–5–4 and 2–1–3–4–5), iii and iv (1–3–4–5–2 and 4–1–2–3–5), and v and vi (1–5–2–3–4 and 2–3–4–1–5). Time for solution does not seem to depend on the anagram as an array with directional properties. This conclusion might not hold for anagrams involving a larger number of letters but its holding for five-letter anagrams accounts for the exceptional findings concerning arrangement IV. In this arrangement (5–3–2–4–1) there is a displaced letter, '4', which need only be moved to the second position to give the solution spelled backward. (3*b*) A shorter solution time is entailed if the displaced letter must be moved a lesser rather than a greater number of places. This follows from the finding that i (1–2–3–5–4) requires a one-place move and a shorter solving time than iii (1–3–4–5–2) which requires a four-place move; and ii (2–1–3–4–5) requires a one-place move and a shorter solving time than iv (4–1–2–3–5) which requires a four-place move. (3*c*) A shorter solution time is entailed if the displaced letter must be moved *from* one of the two extreme positions rather than *from* an intermediate position. In iii (1–3–4–5–2) the displaced item must be moved from the fifth position forward to the second; in v (1–5–2–3–4) from the second position backward to the fifth; and the former has a shorter solution time than the latter. Likewise, in iv (4–1–2–3–5) the displaced item must be moved backward from the first position to the fourth; and iv is solved more quickly than vi (2–3–4–1–5) where the displaced item must be moved forward from the fourth position to the first. An extreme letter is, as it were, more manipulable than an intermediate one.

IV. ADDITIONAL RESULTS

Practice effects

The anagrams taken first, second, third, and so on by one half of the S's are taken last, second last, third last, and so on by the other half. Consider the 36 anagrams taken first by one half of the S's and last by the other half. Of the former, 24 are solved in 10 sec. or less and 12 in 11 sec. or longer (10 sec. is the overall median solving time); of the latter, 22 are solved in 10 sec. or less and 14 in 11 sec. or longer. Applying χ^2, the difference between the two proportions (i.e. 24:12 and 22:14) falls far short of significance. Nor is there any noteworthy difference between the corresponding proportions for: the anagrams taken second and second last (20:16 and 20:16); the third and third last (20:16 and 17:19); the fourth and fourth last (17:19 and 17:19); and the fifth and fifth last (18:18 and 19:17). There are, therefore, no practice effects within the anagram lists. If solving times alter as a result of practice within the session, this is either confined to the four practice anagrams or is masked by some opposing change.

The distribution of solution times

Suppose that a computer is constructed which has the following characteristics: it has a store of all five-letter words, a device which permutes the five presented letters either systematically or at random, and it compares each successive permutation against its store of words until a match is found. If such a computer were given a number of anagrams to solve and if each anagram allowed of only one solution, then the solving times would be normally distributed about a mean representing the time required to perform some sixty permuting and matching operations. Such a distribution might not, however, arise in the case of a computer whose permutations were programmed according to, say, information about the transition probabilities between letters for, with such additional programming, the machine would achieve a number of earlier successes: and the greater the number of relevant rules built into the machine, the faster would it achieve successes.

The second row of Table 1 shows all 864 solution times in a cumulative frequency distribution with 3 sec. intervals. The third row shows each total divided by 72, that is, shows the average number of solutions (to one decimal place) achieved up to any specified time by the average S taking the average list of twelve anagrams. The last row shows the differences between successive averages. It can be seen that the number of solutions approximates to a negatively accelerated increasing function of time. (An end-spurt is suggested which may be due partly to S's renewing their efforts as the time limit approaches and partly to E's allowing a generous rather than a strict time limit.) Notice that if the average S spends a maximum of 3 sec. on each anagram, he solves 3 anagrams out of the 12; and if he extends his maximum time by a further 3 sec., he now solves an additional 1·6 anagrams. But if his maximum time is 12 sec. on each anagram, he must now extend this by a further 9 sec. to solve an additional 1·6 anagrams. And if his maximum time is 15 sec., he must extend this by 12 sec. to solve an additional 1·5 anagrams. The conclusion from Table 1 is that the longer S has spent up to the moment without attaining the solution of an anagram, the longer will he have to spend on the average before he attains solution. Mayzner & Tresselt (1958) note this same temporal phenomenon, point out its similarity to the progressively declining speed of producing words in the continuous controlled association situation, and conclude that 'the results...seem to imply that the production of implicit responses or hypotheses to an anagram problem situation can be described by the same law as that governing the production of explicit responses in verbal recall and word problem situations' (p. 379). In general terms, it might be said that, as the trial-and-check sequence progresses, the thinker becomes less able to initiate novel tries. In part, this may be due to a sort of exhaustion of the supply of tries available to the thinker: in part, it may be due to the progressively increasing difficulty of keeping track on which productions have already been tried out, with the result that 'favoured' tries are repeated. It is evident also that the human anagram solver does not perform like the rule-less permuting and matching machine described above.

Table 1

Time (sec.)	3	6	9	12	15	18	21	24	27	30	30+
Total	214	334	407	468	514	549	580	600	622	648	864
Total ÷ 72	3·0	4·6	5·7	6·5	7·1	7·6	8·1	8·3	8·6	9·0	12·0
Difference		1·6	1·1	0·8	0·6	0·5	0·5	0·2	0·3	0·4	

Comparison of the different solution words

Since practice effects are absent from the results of the experiment, the overall solving time for each of the twelve solution words can be examined by the same procedure used above to analyse the effect on solving time of the different anagram arrangements. When ranked in increasing order of difficulty, the words assume the following order (the median solving time for each is shown in brackets). LUNCH (6), HOUSE (6), CHAIR (7), BRICK (9), WATER (9), DEPTH (9), SCARF (10), ABOUT (11), IMAGE (12), UNCLE (21), OTHER (30+) and RATIO (30+). The significant differences are summarized in the following statements. LUNCH is easier than IMAGE ($P = 0{\cdot}008$) and, *a fortiori*, than UNCLE, OTHER and RATIO.

ABOUT is easier than UNCLE ($P = 0{\cdot}0258$) and, *a fortiori*, than OTHER and RATIO. UNCLE is easier than RATIO ($P = 0{\cdot}0042$). No other differences attain the 0·05 level of probability.

The conclusion would seem to follow that the most difficult words are those which begin with a vowel. Apparently S prefers to try out provisional words or letter combinations which start with a consonant rather than a vowel: such a preference reflects the greater number of words which do, in fact, start with a consonant. The exception to this conclusion is the fact that RATIO is the most difficult word of the twelve—presumably because of its being relatively uncommon.

V. A REPLICATION OF THE EXPERIMENT WITH SCHOOL CHILDREN

This replication was conducted by one of the writer's students, Mr Donald N. Mackay. The subjects were twelve 12-year-old (mean age, 12 years 4 months) and twelve 15-year-old (mean age, 15 years 5 months) Junior Secondary School pupils. The replication differed from the main experiment only in the following respects. Each S was tested individually by the same experimenter (Mr Mackay): the list of practice anagrams was lengthened by adding some easier anagrams: the time limit was 60 sec. per anagram: and in each age group, each anagram list was taken by only one S, always in the forward direction. Apart from the small number of S's, the replication is unsatisfactory because of the large proportion of failures to solve. With the 15-year-olds, 59 of the 144 anagrams are failed; with the 12-year-olds, 88 of the 144 anagrams are failed. The conclusion can, nevertheless, be drawn that the relative ease of solving low as opposed to high scrambled anagrams increases from the students to the 15-year-olds to the 12-year-olds. This conclusion emerges from the following comparisons. Consider each individual S. If he solves six or more of the twelve anagrams within whatever time limit is allowed, take the six anagrams solved in the shortest times and find the proportion of these which are low scrambled. If S solves less than six anagrams, find the proportion of low-scrambled anagrams among those solved. To the nearest percentage, the averages of these proportions are 68 for the students, 76 for the 15-year-olds, and 82 for the 12-year-olds. These proportions are significantly higher for the combined group of 24 pupils than for the group of 72 students. Among the pupils, 8 have a percentage of 67 or under, and 16 a percentage of more than 67: among the students, 45 have a percentage of 67 or under, and 27 a percentage of more than 67. The χ^2 value for the differences between the two groups is 5·2261 which yields a P value of 0·02 (to the nearest second decimal).

VI. THE IMPROVABILITY OF ANAGRAM SOLVING PERFORMANCE

Student subjects are, on the whole, better anagram solvers than 15-year-olds who are, in their turn, better than 12-year-olds. In contrast to this, the students do not, within the test list of twelve anagrams, show any beneficial effects of practice. These two findings raise the question: how does improvement come about in a person's overall anagram solving performance? Notice that this question refers to anagram solving in general and not to the special situation where a person takes a succession of anagrams sharing some common characteristic such as the same pattern of letter disarrangement or the same meaning category of the solution. As he progresses from one such anagram to another, his solving times become, on the whole, shorter because he becomes set to try out those classes of proposal which he has found to be successful, e.g. he searches for animal words. In the present experiment, such recurring characteristics are deliberately excluded so that the person cannot become set for the sort of proposal appropriate to the next anagram to be taken. It is to situations where there is little or no opportunity for acquiring specific sets that the above question is addressed.

In attempting to answer the above question, consideration must first be given to the structure of the class of anagram problem, to the rules to which the class conforms. In the solving of three-term series problems, the logical structure of the problem class provides rules which are learned over months of day-to-day activities and also in taking a number of such problems in close succession (Hunter, 1957*b*). So, too, in the solving of five-term series problems. (In an experiment started by the writer but, for various reasons, abandoned, students took a sequence of twelve problems in each of which there were four two-term premises embodying five letter-terms which were to be written down in rank order from tallest to shortest. One such problem was as follows: '*H* is shorter than *C*: *H* is taller than *S*: *S* is taller than *K*: *W* is shorter than *K*.' Variations in problem structure and letters used were counter-balanced over a group of twenty-four subjects, none of whom made any errors in their solutions. The median solving time for the problem taken first, second, third, and so on, was 26, 25, 17, 18, 16, 13, 12, 12, 12, 11, 9 and 8 sec. respectively.) And it is almost tautologous to say that the same holds for any other problem in which there is a relatively simple logical structure or argument form. In anagram problems, however, the rules are neither few nor simple: they are the numerous complex statistical rules of word structure and letter sequence. Whereas the rules for solving sorites or syllogism problems could be embodied in a relatively short statement, the linguistic rules for the solution of anagram problems would fill a large workbook. The only brief all-purpose advice which could be given to a person about to take an anagram problem would be vague and of the following kind: proceed by trial and check, guiding your tries by what you know of word structures and letter sequences likely to suit the number and selection of letters given.

To return to the main question of this section, the present view of anagram solving as the successive production and assessment of different proposals suggests three components of the activity in which general improvement may take place. The first, and perhaps the chief, component is the thinker's linguistic knowledge. The larger his vocabulary and the more accurate his knowledge of statistical characteristics of words, letter sequences, and speech sounds, the more likely is he not only to attain solution but to do so more rapidly because his earliest proposals have a greater chance of being successful ones. Such linguistic knowledge is, presumably because of its great complexity, characteristically acquired by prolonged commerce with language. This being so, the slow improvement in this knowledge may well account for much of the age differences reported in the experiment. The older subjects would be better anagram solvers than the younger because of their greater linguistic knowledge. And the younger, as contrasted with the older, subjects would find low scrambled anagrams relatively easier to solve than high scrambled because of their poorer linguistic knowledge: the less a person is influenced by linguistic knowledge, the more will his performance be determined by the locations of the letters in the anagram—if, as would seem to be the case, the simpler location rearrangements are the first to be tried out, then a greater proportion of the solutions of the linguistically unskilled thinker will be those which are attainable through these simpler rearrangements. It would also be expected that no increase in linguistic knowledge would take place during the short practice session afforded by a list of only twelve anagrams; consequently, no practice effects would appear.

The second component is the thinker's ability to rearrange the anagram letters. Some rearrangements are relatively simple while others are relatively complex and, especially with anagrams containing many letters, the latter may strain the thinker's capacity for immediate memory. He may 'lose his way' during the rearrangements unless aided either by a systematic plan of action or by a paper and pencil record. In so far as the thinker is impeded by inability to carry out complex rearrangements and in so far as this complexity of rearrangement is limited by his capacity for immediate memory, then, to judge by memory span data, his performance is unlikely to improve over a relatively short practice session. Like the component of linguistic knowledge, the component of complexity of letter rearrangement seems to improve only over time intervals of the order of months or years. Increased ability to carry out complex rearrangements may account for part of the age differences found in anagram solving and, furthermore, could not be expected to enter into a short practice session.

The third component, which must be regarded more tentatively than the other two, concerns the thinker's ability to produce and assess different proposals in rapid succession. The rapidity with which different proposals follow one another may depend both on a sort of fluency factor and on an immediate memory capacity enabling the thinker to keep track of which proposals he has aready tried. The data of the present study do not enable statements to be made about the role of this third component but its potential cannot be overlooked. Thus, Bousfield & Sedgewick (1944) fit curves to their data on continuous controlled recall of words, and calculate two constants, one representing the total supply of words available to the person and the other the rate of depletion of this supply. That these two constants have psychological significance is shown by their assuming different values under different experimental conditions. And again, McNemar (1955) contrasts university students who score high on a test of logical reasoning with students who score low, and observes that 'High and Low Groups do not differ in speed of unrestricted word flow, but do differ in speed of controlled association; the difference, favouring the High Group, increases in significance with each restriction placed on the response' (p. 36).

The question of the improvability of overall anagram solving performance may be answered tentatively as follows. Improvement occurs only as a result of: (1) quantitative increase in linguistic knowledge, making possible an earlier production of proposals which are likely to be successful and also a more likely recall and recognition of appropriate solution words; (2) quantitative increase in immediate memory capacity enabling more complex rearrangements to be achieved and a fuller record kept of proposals already tried; and (3) increase in the speed of producing sequences of potentially appropriate proposals. With the possible exception of the last factor, these changes are of a kind which cannot be brought about within a practice session of the order of minutes but, rather, by prolonged day-to-day opportunity for learning which is of the order of months or years.

VII. THE RELATION OF ANAGRAM SOLVING TO AWARENESS

The retrospective reports of the students typically allude only to fragments of the thinking process and are remarkable less for what they contain than for what they do not. On the positive side, the reports refer to the grouping of two or more letters, to the use of a letter or letter group as a focus round which the remaining letters are placed in successive arrangements, and to the recalling of a succession of words or syllables which somewhat fulfill the requirements of a provisional skeleton of the sought-for solution. The reports also mention that, on attaining or being told the solution, the thinker often realizes that his solving has been impeded by some persistent and inappropriate 'theme'. Examples of such reports are as follows. HOUSE cannot be attained because the word SHOE or the digram SH dominates the proposals: UNCLE is difficult because the letter 'C' is assumed to be 'soft', and IMAGE is difficult because 'G' is assumed to be 'hard': UNCLE is difficult because it is assumed that the solution must start with a consonant: taking the anagrams used by Hunter (1956), a student knows that the solution is the name of a musical instrument but fails to attain BANJO because, despite his owning and regularly playing this instrument, he 'thought only of orchestral instruments'. Such reports indicate that the total thinking sequence divides into subsequences characterized by a common 'theme' or 'plan of action', that is, anagram solving shows clustering effects similar to those found by Bousfield (1953) in continuous controlled association of words and by Duncker (1945) in his studies of problem solving.

Perhaps the most puzzling feature of the retrospective observations concerns the relative distinction between what might be called implicit and explicit activity. Either kind of activity is reported as occurring at any time during the 30 sec., but implicit activity seems characteristic of the initial few seconds while explicit activity typifies the later part of the thinking sequence. Regarding the former, the report is to the effect that 'I didn't do anything but just looked at the letters waiting for the solution to come' and often the solution 'springs out suddenly and spontaneously without my having anything at all to do with it'. In explicit activity, the thinker is aware of 'doing things' although he may find such 'doings' difficult to describe. This distinction between implicit and explicit raises two questions which will serve to end this paper on an admittedly speculative note. (1) What is the thinker doing at those times which, in retrospect, he describes by saying that he is 'doing nothing'? The writer would suggest that he is engaged in activity which is substantially similar to the activity retrospectively described as 'trying things out'. In short, the suggestion is that the thinker is engaged in a trial and check activity which is continuous (except perhaps for the occasional intrusive 'fantasy' not directed to solving the problem in hand). (2) If this is the case, why should some parts of the thinking sequence be available to retrospective report and other parts not? It may be suggested that, in the main, the answer to this complex question is that implicit, as contrasted with explicit, activities are those which are more easily available to the thinker as a result of being well practised, automatized modes of perceiving or using words, letter sequences and speech sounds; implicit activities are modes of organizing and judging the data which are, as a consequence of past experience, achieved 'without thinking'. Three considerations somewhat support this interpretation. First, students never offer any comment concerning the presumably very easily accomplished process of judging the appropriateness of proposals. Secondly, implicit activity typically occurs during the earliest part of anagram solving, that is, at that time when, according to the present view, the thinker tries out his most readily available proposals (when the solution fails 'to appear' during this early period, students often report mild annoyance at having to 'get down to work and do things'). Lastly, in discussing the regrouping of letters, students often speak as though it is the letters themselves which initiate the grouping without the thinker taking any part in the process: some letters 'come together', 'obtrude themselves', and 'attract each other'. It is obvious, however, that letter groupings are achievements of the thinker: the same groupings would not 'obtrude themselves' for the illiterate. Now, it has often been noted that when objects become familiar through learning, their perceived characteristics seem, to the perceiver, to inhere in the objects themselves. Likewise, familiar learned rules of inference appear to inhere in the data of the problem. By analogy, well learned transition probabilities between letters may be attributed, by the anagram solver, to the letters themselves. If it is permissible to turn this supposition around, it may be that the solver's attributing the achievement of a letter grouping to the letters is evidence for this grouping's being an 'old friend', being familiar as a result of having been commonly encountered.

The writer is indebted to Dr Boris Semeonoff for many helpful discussions of statistical matters, to Mr Donald Mackay for replicating the experiment with school children, and to Mr Norman Murchison, O.B.E., headmaster of Ainslie Park Junior Secondary School, Edinburgh, for his kind co-operation in making both pupils and testing facilities available. Although the present paper had been accepted for publication at the time, the writer was fortunate in being able to incorporate the results of the two articles by Drs Mayzner and Tresselt: thanks are due to Dr Mark S. Mayzner for making this possible by providing a pre-publication copy of the 1959 paper and also details concerning digram letter frequencies.

References

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *J. Gen. Psychol.* **49**, 229–40.

Bousfield, W. A., Cohen, B. H. & Silva, J. G. (1956). The extension of Marbe's Law to the recall of stimulus-words. *Amer. J. Psychol.* **69**, 429–33.

Bousfield, W. A. & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *J. Gen. Psychol.* **30**, 149–65.

Bruner, J. S. (1958). Critical notice of *Thinking* by Sir Frederic Bartlett. *Brit. J. Psychol.* **49**, 160–3.

Duncker, K. (1945). *On Problem-solving. Psychol. Monogr.* **58**, no. 270.

Hollingworth, H. L. (1935). The conditions of verbal configuration. *J. Exp. Psychol.* **18**, 299–306.

Hunter, I. M. L. (1956). The influence of mental set on problem solving. *Brit. J. Psychol.* **47**, 63–4.

Hunter, I. M. L. (1957*a*). Note on an atmosphere effect in adult reasoning. *Quart. J. Exp. Psychol.* **9**, 175–6.

Hunter, I. M. L. (1957*b*). The solving of three-term series problems. *Brit. J. Psychol.* **48**, 286–98.

Hunter, I. M. L. (1958). Skill in human reasoning. *Proc. Roy. Phys. Soc. Edinb.* **27**, 33–9.

Hunter, I. M. L. (1960). Remembering. In *Psychology Through Experiment*, ed. by G. Humphrey. London: Methuen. (In the Press.)

Mayzner, M. S. & Tresselt, M. E. (1958). Anagram solution times: a function of letter order and word frequency. *J. Exp. Psychol.* **56**, 376–9.

Mayzner, M. S. & Tresselt, M. E. (1959). Anagram solution times: a function of transition probabilities. *J. Psychol.* **47**, 117–25.

McNemar, O. W. (1955). An attempt to differentiate between individuals with high and low reasoning ability. *Amer. J. Psychol.* **68**, 20–36.

Pratt, F. (1942). *Secret and Urgent: The Story of Codes and Ciphers.* New York: Blue Ribbon Books.

Siegel, S. (1956). *Nonparametric Statistics.* New York: McGraw-Hill.

(*Manuscript received* 6 *February* 1958)

REMEMBERING ADVANCE CUES DURING SEARCHING

By N. H. MACKWORTH and J. F. MACKWORTH

Medical Research Council, Applied Psychology Research Unit, Cambridge

A classification is given of the various ways in which advance information may be displayed to help performance, and a distinction is drawn between two kinds of work. These are tasks in which advance information (1) always remains on view until after action has been taken, and (2) is displayed but must be memorized, because it disappears before action is required.

A particular laboratory task of order search has been studied; this has the advantage that the facts needed for a series of decisions are displayed all the time, and therefore do not need to be remembered. But such a display also has the disadvantage that information needed for use at the moment cannot easily be selected from many other details. Visual searching can be greatly simplified by providing additional advance cues which emphasize in turn each objective as it requires attention. However, if these advance cues point to objectives which are too far ahead, once again difficulties arise in the memorizing of these extra cues.

The experiments described below investigated the size of the *advance store*, which was the number of advance cues which could be memorized at any given instant for successful work; this advance store did not include any advance cues currently on the display.

Studies of the advance store demonstrated that serial work greatly reduced the normal immediate memory span. Only two advance cues could be memorized to help subsequent performance at slow working speeds, and at fast speeds only one such cue could be remembered. The *advance time* was also much less with the faster work.

The limiting factor for anticipation span at the fast speeds was apparently the time available for perception and memorizing. In keeping with this was the further result that grouping the advance cues together in time gave better performance, and some of the possible reasons for this have been discussed.

I. INTRODUCTION

A good deal of work has been carried out on immediate memory, the subject being given a series of items and asked to repeat them. One of the most detailed investigations on the relation between the nature of the material and the memory span was carried out by Brener (1940). Under ideal conditions, such a test gives a memory span of 7 to 12 digits for normal adults. However, as soon as the memory test becomes more complicated the memory span falls abruptly. The experiments described in this paper were designed to see how far ahead advance information could be usefully given in a simple decision-taking task. In other words, the intention was to find out how many items a person could remember at a time in a continuous task, when this store of items was continually changing. Although the store remained the same size throughout the task, he had to discard items at the beginning after use and add items at the end of this store.

II. ANTICIPATION IN THE PERFORMANCE OF TASKS

An important distinction has been drawn by Poulton (1950, 1952*a*, *b*, 1957) between the two main kinds of anticipation found in human behaviour—receptor and perceptual anticipation. Skilled behaviour usually has the characteristic that observation runs ahead of performance; information, therefore, precedes action in both forms of anticipation (Bartlett, 1951*a*, *b*, 1958). The distinction depends on the way in which this information is used.

Receptor anticipation implies only that the subject is shown future stimuli before he has to respond to them, i.e. he is allowed preview by the display. Perceptual anticipation entails predictive behaviour by the person concerned; prediction can be regarded as the forecasting of future events from an estimate of their probabilities derived from existing information—in other words the use of experience from the past to predict the not-quite-yet. This can be done by recalling a similar situation from a previous occasion, or, alternatively, by perceiving a characteristic pattern in space or time which gives a guiding idea for subsequent action. Vince (1955) has made the further point that such predictive interpretation may be based not only on the changing visual signals themselves, but also on the person's own responses to these changes.

Receptor anticipation and preview

This paper is concerned with receptor anticipation, which is a rather simpler process than prediction. Much work has been reported on the temporal relationships between stimulus and response, and these studies can be classified as follows: (*a*) no preview; (*b*) preview without memory; (*c*) preview with memory.

(*a*) *No preview—only the present signal displayed*

The absence of preview means that the subject is prevented from seeing anything ahead of the present signal, i.e. he is not shown a signal until a response to it is needed. Only the present signal is activating the sensorimotor system; no advance information is waiting, either to be perceived in the display, or to be used in the brain. Action follows the signal after the reaction-time (Hick, 1952; Hyman, 1953). Absence of preview is also found in a self-paced situation where response to the present signal releases the signal for the next response, as, for example, in the work of Knowles, Garvey & Newlin (1953).

(*b*) *Preview without memory—the present signal and its successor(s) displayed*

Preview occurs when the subject is allowed advance information during a serial task, so that while he is responding to the present signal he can be considering a subsequent signal. Welford (1952) has discussed various aspects of this response-signal overlap. (Some other but related term such as '*static preview*' is needed for situations in which the subject is allowed a complete preliminary survey of his task; before he starts to respond, he sees the pattern of all the signals he will shortly be using serially.)

In *preview without memory* the subject is not obliged to remember any advance information because this remains on the display until after action has been started. With simple sensorimotor tasks not requiring memory, the present signal should be overlapped by its successor for effective performance (cf. Welford, 1952). No additional advantage is derived from displaying further information, if the subject is not able to group the advance signals together mentally. Indeed, with more complex perceptual tasks, such extra information can become a serious handicap when the present objective has to be located by visual search with much discarding of detail (Mackworth & Mackworth, 1958).

Most of the existing experimental evidence on preview without memory is on simple sensorimotor tasks. Leonard (1953) has, for example, studied the effect of preview on response when subjects had to move either a joystick or a stylus in the direction indicated by a light. Preview was provided by showing the direction of the next movement during the present response. It was found that allowing this response-signal overlap gave definite advantages. Performance was, in general, faster, because the responses started more quickly; these movements were also smoother although slightly slower.

Wagner, Fitts & Noble (1954) asked their subjects to report the transit of dots passing behind a cross-line at the rate of two dots per second. When only the present signal was on view, less than half the responses were made at the correct moment. Accuracy was optimal when the subject could see the present signal and its successor. No further improvement was obtained when the display simultaneously showed four or even twelve signals. Poulton (1954) has reported almost identical results from a display presenting a series of eight-way choices. Performance improved when the following signal was added to the present signal, but further additions to the display gave no further benefit.

(c) *Preview with memory—advance signals or cues being remembered*

Advance signals can take the form of items which are themselves stimuli for later responses. Alternatively, they can be *advance cues*, or changes in the environment which emphasize material, related to a future decision, already showing in the display. *Advance information* is a conveniently brief term for either advance signals or advance cues.

During continuous work, advance information may be displayed and then withdrawn before the situation has developed far enough for action to be undertaken. The subject has then to memorize this information as he goes along, a requirement in the display which is classified in this paper as *preview with memory*.

There are two ways in which the extent of receptor anticipation can be estimated during preview with memory in continuous work. The first is to measure the period during which advance cues are stored before they give rise to action, and this is the *advance time*. In the test described in this paper, this is the number of seconds between the onset of the advance cue and the deadline for action on the particular objective underlined by that cue. The second method is to measure the size of the *advance store*, which is the number of advance cues that have already been presented to the sense organs, but have not yet been acted upon. These advance cues have vanished from the display, and are (or ought to be) in the memory store, ready for use when necessary. The advance store therefore does not include any advance cue at present on the display.

Preview with memory can take at least two forms. These are: (1) without any display grouping of advance information; and (2) when advance information is grouped in time or space.

(1) *Preview with memory—without any display grouping of advance information.* In a continuous multiple-choice situation, Poulton (1954) found that when only one signal was on view at any given moment, performance began to be impaired when that advance signal was further ahead than the signal indicating the present response. When the size of the advance store was varied between zero and three items, the greater the size of the store, the greater the decrement in performance; the decline was statistically significant when two or more signals were being stored. Moreover, only two items could be successfully stored at a slow speed, and at a faster speed only one item could be stored to give a performance significantly better than chance. Thus performance deteriorated despite the fact that at these faster speeds the time during which signals had to be remembered was less. The inference was that adequate time for perception outweighed any decay process (over the range of advance times studied which was between zero and 6 sec.).

Kay (1953) also carried out an experiment requiring subjects to press a key in response to a light. The correct response could be delayed by the experimenter so that it was separated from its stimulus light by a variable number of other stimuli. He found that correct responses fell from 90 % at one back to 50 % at two back, and were down to 15 % by three back. He also found that the memory span measured in this way was less as the speed of presentation was increased.

(2) *Preview with memory—when advance information is grouped in time or space.* When the subject can pay attention to a group of signals at once, and store them before he is required to respond to any of them, then he may be able to do the task more satisfactorily (although not usually so well as he can when he is shown only one signal ahead of the next).

Grouping of items in time has been studied by Knowles & Newlin (1955), who found a maximum of four signals remembered during a serial task. These signals came from a 5 × 5 matrix of display lights, above a similar response panel, and were presented in groups of 2, 3 or 4 separate lights in succession at fixed speeds. After the display of a group of signals, the subject responded to the group at his own rate, and then a new group was presented automatically. The results were compared with performance when the task was self-paced (single lights required a response before the next light was given). Performance was slightly faster with the lights grouped together in two's and given at the highest speed, than with this single-light self-paced situation.

Display grouping of items in space and time has been considered by Vince (1949), in some tapping experiments, for which the signals were dots. Response was much faster in a paced task when columns of these dots appeared simultaneously through a narrow slit, than when the same pattern of signals was presented successively in the form of a comparable broken row of dots. The average rate of response was 5 per sec. when the display consisted of columns of dots, which were thus grouped in both space and time; but this speed fell to as little as 2 taps per sec. with the broken row display, which grouped the material in time only.

III. EXPERIMENTS ON THE EFFECTIVE RANGE OF REMEMBERED ADVANCE CUES

(*a*) *Aims of present research*

The investigations described in this paper consider an example of preview with memory, the general purpose of the work being to estimate anticipation span during a decision-taking task.

The problems investigated in the first experiment were: (1) to determine the anticipation span during continuous decision-taking work at a given speed; and (2) to note whether increasing this demanded speed had any effect upon the span, as measured by the advance time and also by the advance store. In this first experiment the advance information was not grouped in any way; in the second it was grouped in time only.

(*b*) *Experiment* 1. *Preview with memory, without display grouping—a comparison between fast and slow speeds*

(i) *Method*

The basic procedure was the order search task previously described (Mackworth & Mackworth, 1958). Briefly, the display consisted of 20 counters which had to be moved down their respective columns in a grid in time to the ticking of a 3-figure clock. At each tick of the clock, 1 counter could be moved on one square down its column. In each column of 20 squares, 4 squares contained 3-figure numbers. These numbers were regarded as objectives in space and time; they were arranged in such a way that if the counters were moved correctly each one would arrive at each numbered target square at the moment that that number was showing on the clock. The task for the subject was to decide which counter should be moved, by comparing the various times shown on the display. He called out the letter identifying this counter and the experimenter moved it one square down the column. The speed of the clock was set so that at the fast speed it ticked once every 1½ sec., requiring a new numbered square to be found and therefore a new decision to be made, eight times a minute; at the slow speed the clock ticked at half this rate.

The displays had no moves to spare, and therefore the correct counter had to be moved every time the clock ticked to obtain a perfect score. (Minor errors were not, however, included in the final results as discussed below.) The subject was told that it was just as wrong for a counter to arrive too early at the numbered squares as for it to arrive too late. The test was marked by noting the difference between the actual time of arrival of the counter at a numbered square, and the number shown on the square. It was pointed out to the subject that if he left out many moves, he would find himself in an increasingly awkward position owing to the cumulative nature of the process. Thus he was biased to keep the counters moving and to make mistakes when the task became difficult, rather than to leave out some moves altogether.

Table 1. *Example showing relation between spot light position and required response at different store sizes*

Clock times	0	0–3	4–10	11–13	14–18	19–26	27–30	31–33	34–39	40–46
Response (counter to be moved)	—	C	E	A	H	I	K	G	D	B etc.
	Objectives outlined by spotlight in column named									
Store size										
0	—	C	E	A	H	I	K	G	D	B
1	C	E	A	H	I	K	G	D	B	—
2	CE	A	H	I	K	G	D	B	—	—
3	CEA	H	I	K	G	D	B	—	—	—
4	CEAH	I	K	G	D	B	—	—	—	—

The material was presented by means of a film strip, the visual size of the display being as before. Three different films were used, so that each of the twelve subjects did each of these twice at each working speed. The display consisted of white numbers on a black ground. In each picture, one objective was outlined by a white square, successive objectives being spotlit in this way in successive pictures.

The sudden additional emphasis given by the arrival of the spotlight on one particular objective in the display can be regarded as an *advance cue*. It took about half a second to change the picture, and this was done just before an objective time appeared on the clock. As that objective time came up, the spotlight was in its new position so that when the subject sought his new objective, the new situation was waiting for him.

It was thus possible to indicate a future objective at any stage ahead of the present one. The subject had a constant number of objectives to remember for any given test run; a constant store size, as it were. Table 1 shows the relationship between the objective being spotlit and the required counter movement at that particular stage of the test.

With a *store size of* 0 the subject had nothing to remember because the spotlight was always pointing to the *present objective*, i.e. the one towards which he should be moving now.

With a *store size of* 1 the objective emphasized was always the one after the present one, i.e. the objective next-but-*one*. Here the subject had to perceive and remember *one* position of the spotlight while he was moving his counter towards the present objective. (Just before the counter reached the present objective, the spotlight left the new objective which should have been memorized by then.) With a *store of* 2 the objective emphasized was always that next-but-*two*, and *two* objectives had to be memorized. A *store of* 3 meant that *three* objectives had to be retained. Similarly, a *store of* 4 meant *four* objectives had to be remembered.

A control condition, termed *unaided search*, was also studied, and during these runs the spotlight was providing useless information, since it emphasized a past objective which was six earlier than the present one.

Experimental design. Twelve naval ratings, aged between eighteen and twenty-three, acted as subjects, and each man did all twelve conditions; six sizes of advance store being given at each of the two speeds. These store sizes were presented in a randomized Latin square arrangement, and at each session the subject did both speeds at the same store size. Half of the subjects did one alternating scheme of speeds, and the other half had the opposite arrangement. They practised at both speeds for an hour in the first session on the displays, with stores of 0 and 1 and with unaided search. Then, at each session the subject was given a practice run on the store size for that session, at both speeds.

As the subject reached each of his objectives, the experimenter read out the expected and actual times of arrival, and these were recorded. Thus the subject's own performance was emphasized to him as he went along, information which was also available in his display.

Scoring. The test was scored by counting those objectives which were achieved within ± 5 ticks of the correct time. The great majority of these were either absolutely correct, or just after the correct time. There was a tendency for an occasional move to get lost at the fast speed, through no fault of the subject, but the method of scoring allowed for this. A record was also kept of the number of objectives not reached at all.

(ii) *Results from Experiment* 1

Fig. 1 and Table 2 give the results in terms of percentage of correct decisions, based on the number of objectives reached within 5 ticks of the expected time of arrival.

An analysis of variance was carried out on these results. This is shown in Table 3. It shows significant differences due to subjects, speed and advance store sizes ($P < 0{\cdot}01$).

The advantages of the advance store sizes of either 0 or 1 are clear from Fig. 1. The use of the spotlight on the present message reduced the possible confusion between the

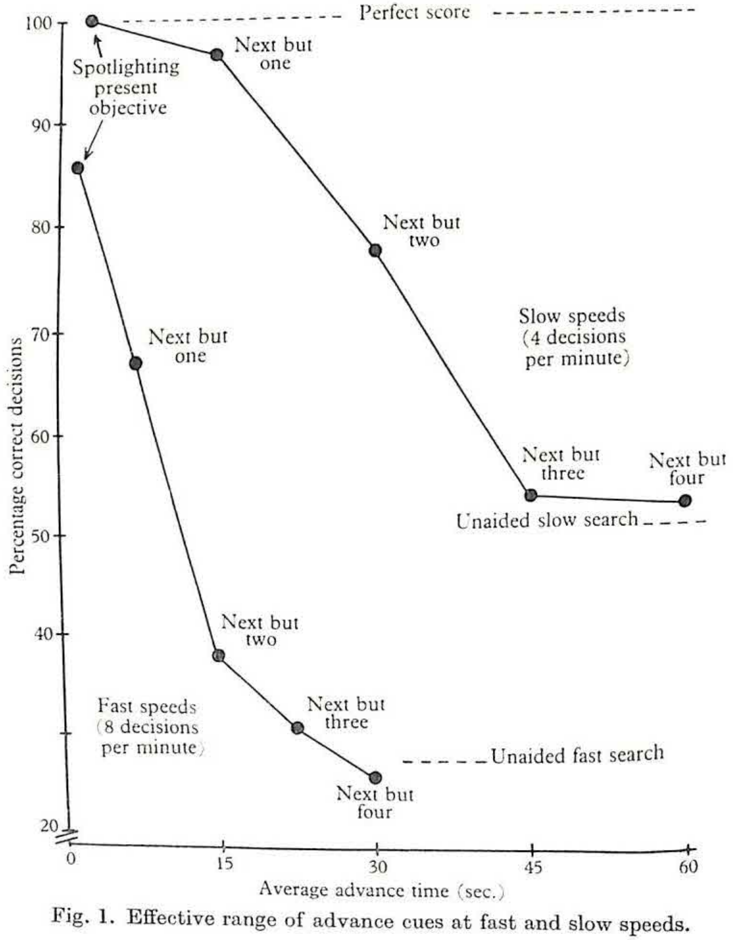


Fig. 1. Effective range of advance cues at fast and slow speeds.

wanted and unwanted objectives to an extent that virtually eliminated visual searching, especially at the slow speed, i.e. a store of 0 or 1 gave 100% or 97% success (see Table 2). At either speed, store sizes of 0 or 1 gave results which were significantly better ($P < 0{\cdot}01$) than those at any of the other conditions which tried to place greater burdens on the memory.

Table 2. *Percentage correct decisions in relation to stage of advance information in Exp.* 1

	Advance store size					Unaided search
Speed	0	1	2	3	4	
Fast	86	67	38	31	26	27
Slow	100	97	77	53	52	50

Table 3. *Analysis of variance for Exp.* 1

Source	D.F.	S.O.S.	M.S.	V.R.	Significance
Differences between grouped pairs of subjects (G)	5	962	192	1·3	N.S.
Advance store size (A)	5	42,561	8,512	**59·1**	< 0·01
Practice between successive pairs of runs (P)	5	2,696	539	**3·7**	< 0·05
Residual 1	20	2,872	144	—	—
Latin square blocks	35	49,091	—	—	—
Speed (S)	1	14,864	14,864	**228·7**	< 0·01
Order within pairs of runs (O)	1	21	21	0·3	N.S.
Differences within pairs of subjects (D)	6	2,547	424	**6·5**	< 0·01
$O \times S$	1	31	31	0·5	N.S.
$O \times G$	5	170	34	0·5	N.S.
$S \times G$	5	227	45	0·7	N.S.
$A \times S$	5	1,306	261	**4·0**	< 0·05
$A \times O$	5	299	60	0·9	N.S.
$P \times S$	5	106	21	0·3	N.S.
$P \times O$	5	289	58	0·9	N.S.
$\Sigma(D_i \times O)$*	6	1,001	167	**2·6**	< 0·05
$\Sigma(D_i \times S)$*	6	1,940	323	**5·0**	< 0·01
$O \times S \times G$	5	535	107	1·6	N.S.
Residual 2	52	3,374	65	—	—
Total	143	75,801	—	—	—

* These terms represent the sum of the interactions of the pairs of subjects with order or with speed.

Black figures indicate a statistically definite effect from that source.

The analysis of variance also showed a significant effect of the interaction between the size of the advance store and the speed of work ($P < 0{\cdot}05$). The nature of this relationship is demonstrated in Fig. 1 and Table 2 when one compares the effects of a store size of 2, at the slow speed and at the fast speed. At the *slow speed* this store size of 2 definitely helped performance because the reading of 77% correct was significantly better than the 50% reading for the unaided slow search ($P < 0{\cdot}01$); it was also significantly better than the 53% and 52% readings for stores of 3 and 4 respectively ($P < 0{\cdot}01$). But at the *fast speed* this store of 2 did not help performance since the reading of 38% correct was *not* significantly better than either the 27% correct reading for unaided fast search or the 31% or 26% readings for store sizes of 3 or 4 respectively.

Fig. 2 also emphasizes the benefits found when people were given two objectives to remember at slow speeds—advantages which disappear when the same people were given two objectives to remember at fast speeds, the evidence now being in the form of the individual scores. It will be seen that with the advance store of 2 at the *slow speed* as many as eight of the twelve men achieved a ‘good score’ of more than 70% correct, but with unaided slow search only one of them reached this level. At the *fast speed* with a store of 2, only one of the men reached this standard and none did with the unaided fast search.

The results in Fig. 1 can also be considered from the point of view of the maximum effective range of advance cues in terms of time, i.e. the maximum *advance time*, the greatest number of seconds during which the advance cues could usually be remembered for use. The differences between the results of the two working speeds were even greater when they were measured in this way. At the *slow speed* the spotlight was still effective even when it was on the objective next-but-two, which represented a maximal effective advance time of about 30 sec. But at the *fast speed* the furthest ahead that benefit was received was with the objective next-but-one, which represented a maximal advance time of about only 7½ sec.—one-quarter of the period for the slow speed.

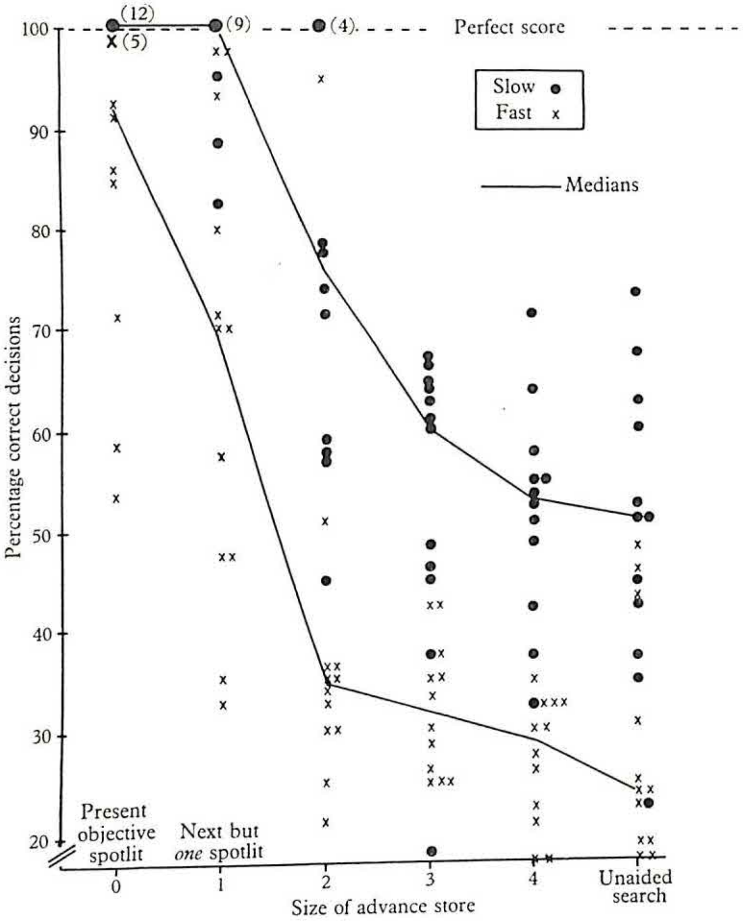


Fig. 2. Individual scores on advance cues at fast and slow speeds.

Fig. 3 compares the percentage omissions at the fast and slow speeds. These non-arrivals were relatively few compared with the arrivals at the wrong time. Store sizes of 0 and 1 benefited performance by reducing omissions at either working speed. The store size of 2 had, however, no effect on omissions at either speed—a point of interest because the advantages seen already with a store size of 2 at the slow speeds must therefore have been due to improved accuracy of timing, rather than to fewer omissions.

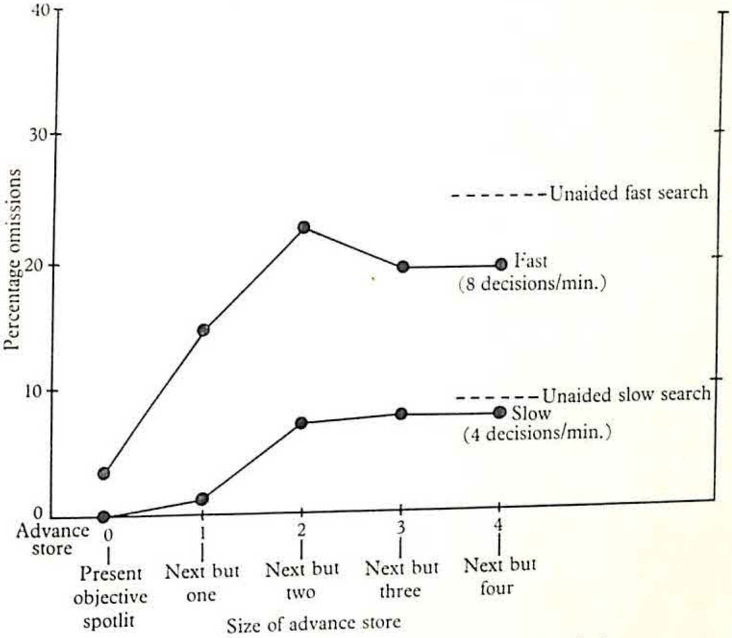


Fig. 3. Effective range of advance cues in terms of omissions.

(c) *Experiment* 2. *Preview with memory: the effects of display grouping*

(i) *Introduction*

The above evidence suggested that two advance cues were the most that could usually be remembered when each of these was given while the subject was dealing with a previous objective. It seemed much harder to perceive, retain and recall items which were constantly changing during a serial task, than to do this with the same items in the classical memory situation.

Two possible reasons for such a short span in serial tasks were considered in planning Exp. 2. The first was the question of the *turnover in the store*. For each new objective that was spotlit, the subject had to omit one past objective and to add a new one to his store of moves yet to be made. This continual changing and reconstruction of the store could have made the task more difficult, particularly when it was so important to recall the objectives in the correct sequence. Exp. 2 was therefore designed to study this factor of turnover; consequently, displays which gave successive objectives in groups were compared with those which provided the objectives singly. (The term 'group' is being used in Exp. 2 for a succession of items of information not separated by responses.)

The second possibility was that in Exp. 1 the memory span was short because of considerable *difficulties in perception and memorization* (see subsequent Discussion). Such difficulties were therefore minimized in Exp. 2. Throughout this experiment the investigator emphasized future objectives by naming aloud the appropriate counters phonetically, and so alternation of visual attention between cue and present objective was avoided. Also the subject no longer had to respond while he was being given the advance cues, because a pause in required response was given in which to listen to the cue.

(ii) *Method*

Display cards with numbered objectives to be reached by the moving counters were prepared as before; each card had twenty columns each of which showed four objectives. Two sets of these cards were constructed, one for the grouped advance cues, and the other for the single advance cues.

In the *grouped task*, the advance cues were presented in groups of either four or five at a time, the two conditions being identified by the terms '4-grouped' or '5-grouped'. The 4-*grouped* task was as follows. Before the clock was started, the subject was told in the correct order the letters indicating the first 4 counters that had to be moved. Then the clock was switched on, and ticked at the fast rate of $1\frac{1}{2}$ sec. per tick. The subject had to reach these 4 objectives in succession, without any moves to spare. No action was required for a further 4 ticks, while the subject was listening to the letters identifying the next 4 counters to be moved. For the 5-*grouped* task the arrangements were the same, except that there were now 5 successive objectives to be reached without a pause, followed by a period of 5 free ticks while the next 5 counters were named.

In the *grouped tasks* each letter always had to be remembered while 3 or 4 other letters were being received or used. The store size in the 4-grouped case varied from 4 letters at the beginning to none at the end. Similarly, the advance time for the 4-grouped task varied from 4 ticks to 20, i.e. 6–30 sec.

In the *single tasks*, the advance cues were presented verbally one at a time. Immediately after the clock reached the expected time of arrival for an objective, there was a pause lasting 1 tick, during which a future counter was named. Apart from this presentation during a pause, the single task resembled the conditions of Exp. 1 with advance store sizes of either 2 or 3, when emphasis was always on the objective next-but-2 or next-but-3. The advance times were between about 15 and $22\frac{1}{2}$ sec.

One move was allowed per tick, and 400 moves were always required for each display. With one extra tick per objective, the total of 480 ticks per run thus lasted for 12 min., which represented 6·7 decisions per minute.

These pauses made the task slightly easier than at the fast speed of Exp. 1, for they allowed the subject to catch up the occasional move inadvertently lost. For this reason, the task in Exp. 2 was marked to rather a stricter standard, an objective being regarded as achieved only if the counter actually arrived there within ± 2 ticks of the expected time of arrival.

Ten naval ratings took part in Exp. 2. Each man received the same design in four separate sessions, held during the morning and afternoon of each of two successive days. A 1 hr. practice session was given on a previous day.

The design used was as follows:

	Morning	Afternoon
First day	2 single, 4 grouped	5 grouped, 3 single
Second day	4 grouped, 2 single	3 single, 5 grouped

Owing to the range of ability between individuals, it was necessary to have two levels of difficulty in both the grouped and the single displays, but these levels of difficulty for each type of display were pooled for statistical analysis.

(iii) *Results from Experiment* 2

It can be seen from Fig. 4 and Table 4 that the grouped cues led to better average scores than did the single cues at the same stage of practice, and this difference was significant, $P < 0{\cdot}01$.

The experimental design did not permit other comparisons, but the following point should perhaps be mentioned, although it cannot be statistically validated. It would appear from Fig. 4 and Table 4 that the 2-*single* display, with a store size of only 2, was about as difficult as the 5-*grouped* condition.

Nine subjects out of the ten made a better score on the grouped than on the single displays, and all of them preferred the grouped displays.

Analysis of the trend of performance during the 12 min. test showed that the 4-*grouped* cues led to an average performance of almost 100 % correct decisions with the first group of 4 letters, but this deteriorated to 70 % by the fifth group of letters, and there was little

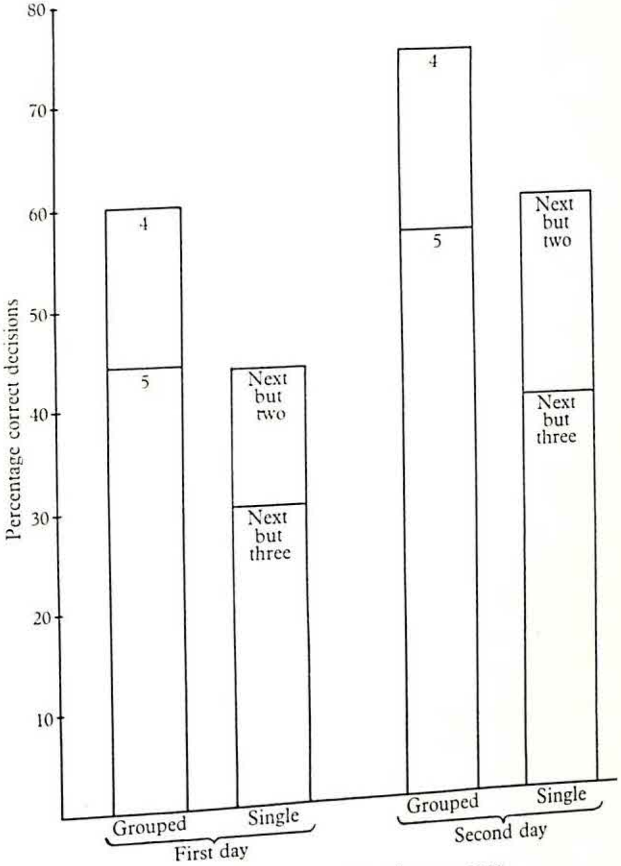


Fig. 4. Effects of grouping advance cues.

Table 4. *Percentage correct decisions in Exp. 2*

	Grouped		Single	
Day	4-grouped	5-grouped	Next-but-two	Next-but-three
First	60	44	43	30
Second	75	57	60	40
Total	67	50	51	35

Table 5. *Sequential analysis of initial performance—percentage correct decisions in Exp. 2*

A. Successive objectives

Display condition	1–4	5–8	9–12	13–16	17–20
4-grouped	96	89	92	86	72

Display condition	1–5	6–10	11–15	16–20
5-grouped	78	75	70	64

Display condition	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2-single	100	100	95	90	70	80	75	85	75	75	75	75	55	65	70	70	55	65	60	65
3-single	100	100	85	50	70	45	50	35	40	60	45	35	45	40	45	45	15	55	20	45

B. Position of objective within group (1st half of run)

	1st	2nd	3rd	4th	5th
4-grouped	94	75	67	73	—
5-grouped	72	56	40	36	38

change after this. With the *single advance cues*, there was an even more rapid initial decline in achievement, as can be seen in Table 5, on the average results from the first 20 objectives with the four display conditions. The first response in a group was always better than the others, but with the 4-grouped there was little difference between the other three responses. With the 5-grouped, the second response was better than the last three.

IV. Discussion

(*a*) *Experiment* 1: *at fast and slow speeds*

The order search test provided a situation where there was complete or static preview of the whole task, but the very complexity of the display led to some difficulty. The addition of the spotlight structured this display, so that a different kind of preview situation was created, one into which the factor of memory could be introduced. Attention had to be given to forthcoming objectives now being spotlit, while responses were being made to achieve the present objective. In Exp. 1 there was preview with memory, but without any display grouping of the advance cues. Maximum performance was obtained when the present objective was spotlit and there was an advance store of 0, the subject not being required to remember anything. Performance deteriorated when the advance store was increased beyond a certain size; it then became too difficult for the subject to remember the advance cues. He therefore tended to search for the required present objective and to ignore the advance cues.

It can be seen that the subjects could obtain help from remembering not more than two cues at the slow speed and one at the fast speed. These figures are surprisingly low when compared with the normal memory span of about seven items (Miller, 1956). Thus a serial task, with its continuous activity and turnover in the signals being stored, is a much more difficult form of work than the simple memory span experiment, where only one static set of items is given before recall is required.

The results of Exp. 1 show that in this paced serial task involving simple decision taking, people shortened their anticipation span when they were required to work at a faster pace. Not only was this span much reduced in terms of time, so that the maximal effective *advance time* was shorter, but the maximal effective *advance store* was also reduced, since fewer cues were retained. These results were in agreement with those of Kay (1953) and Poulton (1954).

In other words, the most important factor in the difference between the required working speeds was the *time available for perceiving and memorizing the signals*. The periods of time involved were not apparently long enough to allow any appreciable decay in stored material. Had there been any appreciable decay with time, the shorter retention time entailed when speeds were faster might have been expected to have improved the results at the faster speed for any given store size. For periods of the order of 1 min., the number of items retained was independent of the time during which they had to be retained, but this number depended on the amount of time available for perception.

This arrangement whereby the extent of the advance store and advance time was imposed by the conditions of work can be contrasted with other situations where the subject may either look ahead of the present signal or work behind it. He can then adjust the advance store and advance time to his own satisfaction, as, for example, in copy-typing, reading aloud or receiving morse code. He may tend to increase his eye-voice span on easy material and reduce it when meeting a difficult word (Miller, 1951). But, in general, with freer working situations of this kind, it has been established quite definitely by experiment that very highly skilled people will not only work faster than less gifted performers, but they will also store *more* signals while they are achieving these faster speeds (Bryan & Harter, 1897, 1899).

There is, however, no inherent contradiction between these earlier results and the present findings in which faster speeds meant *fewer* signals stored. The faster speeds achieved by more skilled men are not necessarily attained in the same way as the faster speeds resulting from quicker pacing of the same person. Practice, leading to improved skill, could, for example, produce a greater relative improvement in either perception or storage than in response, and this could build up the increased number of signals stored between signal and response that is so characteristic of skilled performance.

The advance cues were thought to have been of value in the present studies because they almost eliminated the need for visual search. Direct eye movement studies are needed to check this suggestion which is partly based on the subjects' own reports on their approach to the problem. They usually decided early in any given run whether or not they could make use of the advance cues.

The interpolated activity between perception and response undoubtedly affected immediate memory for advance cues (Brown, 1955, 1958). But this could not have been responsible for the difference obtained between the results at the fast and at the slow speeds, because the *required* interpolated activity was exactly the same in both cases. Then, again, at the faster speeds, which showed less storage, the *actual* interpolated activity was less; more responses were omitted at the faster speeds because more objectives were entirely neglected (see Fig. 3).

(*b*) *Experiment* 2: *the effects of display grouping*

Exp. 2 demonstrated that anticipation span in a serial task could be increased by grouping the advance cues together in time, and this display was unanimously preferred. Three possible hypotheses that could account for the advantage of grouped displays are given below.

(i) *Item and pattern storage*

Storage of a group, which may be dealt with as a whole and then discarded, is probably easier than storage of a continually changing pattern of signals, in which any single signal must change its position in the store consecutively from last to first and then be discarded. Moreover, there is a greater chance of interference with the single item patterns since each new pattern is so like its predecessor (Brown, 1955).

(ii) *Separation of action from perception and storage*

The alternation of attention, from perception and storage to response, is probably more difficult in the single display, because such alternation is more frequent with single than with grouped displays. Moreover, although all displays allowed pauses for perception and storage (as distinct from action) the subjects did not always use this opportunity with the single displays. These pauses were more clearly defined with the grouped cues, because the subject was reminded by the first letter in each group that he had to listen for the other letters. Further evidence of the importance of the separation of response from storage in these complex perceptual tasks is to be found in the fact that on the rare occasions when subjects disregarded their instructions and kept answering during the grouped pause (instead of only perceiving and storing), their subsequent use of the information given during that pause was very poor. However, it should be noted that performance on the first item of a group was markedly better than on the rest of the group, even though maximum interference might be expected on this item. Broadbent (1952) and Poulton (1955) have considered the difficulties of simultaneous listening and speaking, and the related problem of memorization during recall has also been studied by Poulton (1953).

(iii) *Retention time and interpolated activity*

Any interpolated activity which prevents rehearsal will greatly reduce the memory span, even when the signals have to be remembered for only a few seconds (Brown, 1955). The effect of this factor cannot readily be analysed in these experiments, but it was probably important. For example, each cue in the 2-*single condition* has to be remembered for a period during which the following 2 signals are being perceived and the 2 previous signals are being answered. In the 5-*grouped condition* each signal has to be remembered for a period during which 4 other signals are either being received or answered. Provided therefore that receiving and answering a signal can be regarded as similar, then the amount of interpolated activity in terms of number of signals was the same for the 2-single and 5-grouped displays, i.e. 4 signals in each case. It is interesting that performance appears to have been approximately equivalent for these two conditions.

V. CONCLUSIONS

1. This study considers a form of receptor anticipation in which future events are identified on the display, and then this identification is removed so that the advance cues have to be remembered. The extent to which advance cues are being retained can be measured by estimating either (1) the *advance time*, which is the time from the start of the advance cue to the start of the corresponding action, or (2) the *advance store*, which is the number of advance cues that need to be memorized, excluding any at present on the display.

2. A procedure, known as the *advance cue method*, has been devised to study the limits of remembering for use during a form of serial decision-taking at speed. This method has examined these limits of the effective range of advance time and store, as well as their optimum values. The subject could do this work either by memorizing the advance cues or by ignoring these and searching the display.

3. Under such circumstances, it was found that the optimum arrangement was to emphasize the present objective. Indeed, a perfect score could then be achieved, provided the demanded speed was not too high. The reason for this improvement was that emphasis by the advance cue greatly reduced the possibility of confusion between the present objective and the numerous other objectives also on the display, but not required at that particular moment. Visual search was thereby almost eliminated.

4. The limit to the size of the largest effective advance store was reached when only two advance cues were being retained and this was very much smaller than the classic memory span reported with single groups of items. This material could be remembered for a maximum of 30 sec. during continuous work *at the slow speed*.

5. When the subjects had to work faster, all they could remember was one advance cue, which was being retained for as little as 7 or 8 sec. These changes *with faster speeds* were thought to have arisen from a lack of time for perception and memorization.

6. A study was made of one of the possible reasons that could have accounted for the limited storage of advance cues during serial work. It has been definitely established that performance improved when advance cues were heard in groups rather than singly, but further work is needed to determine more exactly the reasons for this advantage..

We are greatly indebted to Sir Frederic Bartlett, F.R.S., for many helpful discussions of this work. We also would like to acknowledge the great value of the comments received from our colleagues and in particular the statistical advice and help obtained from Dr Mervyn Stone.

References

Bartlett, F. C. (1951*a*). Anticipation in human performance. In *Essays in Psychology, dedicated to David Katz*, pp. 1–17. Almquist and Wiksells.

Bartlett, F. C. (1951*b*). The bearing of experimental psychology upon human skilled performance. *Brit. J. Industr. Med.* **8**, 209–17.

Bartlett, F. C. (1958). *Thinking*. London: Allen and Unwin.

Brener, L. R. (1940). An experimental investigation of memory span. *J. Exp. Psychol.* **26**, 467–82.

Broadbent, D. E. (1952). Speaking and listening simultaneously. *J. Exp. Psychol.* **43**, 267–73.

Brown, J. (1955). Immediate Memory. Unpublished Ph.D. Thesis. Cambridge University Library.

Brown, J. (1958). Some tests on the decay theory of immediate memory. *Quart. J. Exp. Psychol.* **10**, 12–21.

Bryan, W. L. & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychol. Rev.* **4**, 27–53.

Bryan, W. L. & Harter, N. (1899). Studies in the telegraphic language. The acquisition of a hierarchy of habits. *Psychol. Rev.* **6**, 354–75.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. Exp. Psychol.* **4**, 11–26.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* **45**, 188–96.

Knowles, W. B., Garvey, W. D. & Newlin, E. P. (1953). The effect of speed and load on display-control relationships. *J. Exp. Psychol.* **46**, 65–75.

Knowles, W. B. & Newlin, E. P. (1955). Coding by groups as a mode of stimulus presentation. Naval Research Laboratory Report No. 4604; Washington, D.C.

Kay, H. (1953). Experimental Studies of Adult Learning. Unpublished Ph.D. Thesis. Cambridge University Library.

Leonard, J. A. (1953). Advance information in sensorimotor skills. *Quart. J. Exp. Psychol.* **5**, 141–9.

Mackworth, N. H. & Mackworth, J. F. (1958). Visual search for successive decisions. *Brit. J. Psychol.* **49**, 210–21.

Miller, G. A. (1951). *Language and Communication*. New York: McGraw-Hill.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81–97.

Poulton, E. C. (1950). Perceptual anticipation and reaction time. *Quart. J. Exp. Psychol.* **2**, 99–112.

Poulton, E. C. (1952*a*). Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.* **43**, 222–9.

Poulton, E. C. (1952*b*). The basis of perceptual anticipation in tracking. *Brit. J. Psychol.* **43**, 295–302.

Poulton, E. C. (1953). Memorization during recall. *Brit. J. Psychol.* **44**, 173–6.

Poulton, E. C. (1954). Eye-hand span in simple serial tasks. *J. Exp. Psychol.* **47**, 403–10.

Poulton, E. C. (1955). Simultaneous and alternate listening and speaking. *J. Acoust. Soc. Amer.* **27**, 1204–7.

Poulton, E. C. (1957). On prediction in skilled movements. *Psychol. Bull.* **54**, 467–78.

Vince, M. A. (1949). Rapid response sequences and the psychological refractory period. *Brit. J. Psychol.* **40**, 23–40.

Vince, M. A. (1955). The relation between hand movements and intellectual activity in a skilled task. *Quart. J. Exp. Psychol.* **7**, 82–90.

Wagner, R. C., Fitts, P. M. & Noble, M. E. (1954). Preliminary investigation of speed and load as dimensions of psychomotor tasks. *Research Bulletin Air Force Personnel and Training Research Center*, TR-54-45, pp. 19.

Welford, A. T. (1952). The psychological refractory period and the timing of high-speed performance—a review and a theory. *Brit. J. Psychol.* **43**, 2–19.

(*Manuscript received* 15 *March* 1958)

HUE DIFFERENCES AND BRIGHTNESS DIFFERENCES AS DETERMINANTS OF FIGURAL AFTER-EFFECTS

By R. H. DAY

Department of Psychology, University of Sydney

Earlier published studies of the figural after-effect using stimulus patterns in which there was no difference in brightness between the inspection figure and its background report inconsistent results. The possibility of more than one type of visual after-effect following protracted fixation of a figure is advanced as an explanation of this inconsistency. In the three experiments described the frequency of occurrence of the figural after-effect is investigated using inspection patterns with and without a difference in brightness between the figure and its background. Different stimulus patterns are used to allow for the occurrence of one or more possible types of after-effect. The results from the three experiments clearly show that, with the brightness of inspection figure and ground very nearly equal, the figural after-effect occurs with as great a frequency as in the case of a marked brightness difference. The results are discussed along with findings from earlier experiments and theoretical implications are considered. The general conclusion is that the necessary condition for the occurrence of the figural after-effect is a visible inspection contour constituted of a difference either in brightness or in hue. A brightness difference between the inspection figure and its ground is not essential for the occurrence of the figural after-effect.

I. Introduction

It is now well established that prolonged visual fixation of one stimulus pattern results in phenomenal changes in a subsequently observed pattern. These changes, which include displacements, distortion and variations in brightness and size, have been called figural after-effects. Although numerous variables affecting these visual after-effects have been systematically studied, it has not yet been clearly established whether figural after-effects occur following the fixation of a stimulus pattern in which no difference in brightness exists between the figure and its ground. In the development of a neurophysiological theory of figural after-effects it is of importance to establish whether the phenomena follow the inspection of such a pattern or not.

Earlier studies (Day, 1957, 1958) have demonstrated that the after-effect of seen movement occurs when all parts of the stimulus pattern are equal in brightness, differing in hue only. The experiments to be described arose out of a consideration of the possible similarity of the processes and mechanisms underlying the figural after-effect and the after-effect of seen movement. The occurrence of both phenomena in the absence of a brightness difference within the stimulus pattern would (although by no means confirming) at least offer some grounds for the view that the two phenomena belong to the same class.

This paper describes three experiments designed to establish whether or not the figural after-effect occurs in the absence of a brightness difference between figure and ground. The abbreviations 'FAE', 'IF' and 'TF' will be used throughout to designate 'figural after-effect', 'inspection figure' and 'test figure', respectively.

In reply to a question following a paper to the Hixon Symposium, Köhler (1951) reported that the FAE could occur when there was no brightness difference between the IF and its ground. This condition was achieved by using different hues of equal brightness for figure and ground. Unfortunately, details of this experiment are no longer available, but Köhler (private communication, 1957) has indicated that the IF consisted of two oppositely slanted lines similar to those used in Exp. 2 (see Fig. 1). The effect of fixation of this figure was a displacement from parallelism of two subsequently observed parallel lines.

Hochberg & Triebel (1955) have reported results which are contrary to those obtained by Köhler. The IF was a red annulus on a grey ground of nearly equal brightness. The expected FAE was a shrinkage of a slightly smaller annulus. Since shrinkage did not occur, Hochberg & Triebel concluded that: 'when figures differ from their ground largely (or only) in chroma, no measurable figural after-effects are obtained'. (These writers use the term 'chroma' to mean hue, whereas, in the Munsell system of colour classification, 'chroma' refers to the saturation of a colour.)

A relevant consideration is the possible occurrence, after prolonged visual stimulation, of more than one type of after-effect. Fox (1951) has adduced evidence in favour of such a view. It is suggested that, on the one hand, there occurs an after-effect due to visual adaptation to a spatial norm in a manner described by Gibson (1933) and, on the other, an after-effect arising from some more purely figural process described by Köhler (1940) and elaborated by Köhler & Wallach (1944). As will be seen, this latter type of after-effect is, in some instances, not easily conceived as an adaptation to a spatial norm. The after-effect obtained by Köhler (1951) using an IF of similar brightness to its ground can be attributed to a process of adaptation to a spatial norm of verticality, which process would presumably operate irrespective of the presence of brightness differences providing a visible contour. No such adaptation would be expected in the case of the stimulus patterns used by Hochberg & Triebel (1955) since the after-effect in this instance involved a change in size. Whereas it is reasonable to suppose that spatial norms exist for the principal spatial dimensions such as verticality, horizontality, etc., it is difficult to conceive of any such norms for size. Thus, the contradictory results outlined above may have been due to the fact that different types of after-effect were studied in the two experiments. Such an argument, however, is not supported by the data to be reported below.

II. EXPERIMENT 1

The purpose of this experiment was to compare the frequency of occurrence of the FAE under conditions of marked brightness difference between IF and ground with the frequency occurring under conditions of nearly equal brightness. The IF and TF used (Fig. 1*a*) had been used in earlier experiments (Pollack, 1958) and were known to give a clear after-effect.

Inspection patterns. In order to construct an IF of a brightness equal to that of its ground, drawing paper was painted with standard Ostwald colours to match Munsell hues. It is necessary here to mention briefly the Munsell system of colour specification (Munsell Color Co., 1950). The specification of a particular colour is given in terms of its position on the colour circle (hue), its brightness (value) and saturation (chroma). Thus the Munsell notation 5·0 YR 5/6 refers to a particular yellow-red (5·0) of value 5 and chroma 6. The value scale extends from 0 (black) to 10 (white). On the chroma scale, 0 refers to neutral grey. In order to establish the values of the colours used, their luminous reflectances were measured in standard daylight (illuminant C) and these reflectances converted by means of tables (Nickerson, 1948) to Munsell values. The pattern with the IF and its ground nearly equal will be referred to throughout as HD (hue difference), and the pattern involving brightness difference as BD. The colours selected for IF and ground in the HD pattern were 5·0 YR 5·14/6 and 10·0 BG 5·24/6, respectively. The difference in brightness between these two colours was then 0·10 of a Munsell value unit. The rectangular IF (Fig. 1*a*) was cut from the yellow-red paper and mounted on the blue-green ground 10 in. by 8 in. The $\frac{1}{8}$ in. diameter fixation point was cut from the same paper as the IF and mounted at the centre of the ground. For the BD pattern, the IF and fixation point were the same as those for the HD pattern but were mounted on a white ground. The luminous reflectance of the white was not measured since it was clear that it was markedly different in brightness from the IF.

Test pattern. The TF and fixation point were cut from the same paper as the IF and mounted on a white ground of similar dimensions to those of the IF ground. The IF and TF drawn to full scale are shown in Fig. 1*a* in the relative positions occupied successively in the experiment.

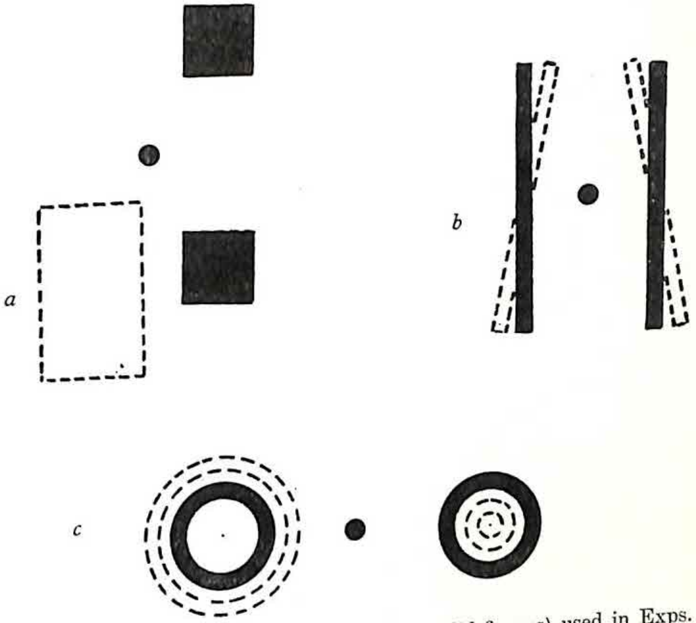


Fig. 1. Inspection figures (dotted outlines) and test figures (solid figures) used in Exps. 1 (diagram *a*), 2 (*b*) and 3 (*c*). In Exp. 1 a figural after-effect was indicated by a displacement to the right of the lower solid square, in Exp. 2 the solid lines were seen as divergent at the top or convergent at the bottom, and in Exp. 3 the right test annulus appeared larger than the left. Both inspection and test figures were solid, but the former are shown in dotted outline for convenience.

The test pattern was mounted in a frame and the IF similarly mounted immediately in front of it so that the two fixation points were superimposed. The frame containing the inspection pattern was hinged at its lower edge so that it could be swung down to expose the test pattern. The distance between the stimulus pattern and the subject's eye was 6 ft. The distance ($\frac{1}{4}$ in.) between adjacent contours of IF and TF subtended a visual angle of approximately 14 minutes.

Subjects. There were 36 subjects, 18 females and 18 males. Ages ranged from 17 to 36 years. The group comprised undergraduates and members of the University academic and research staff.

Design and procedure. In order to obtain viewing conditions as similar as possible for the two conditions of brightness difference, each subject viewed under both conditions. One half of the subjects observed in the order HD, BD, and the remaining half in the opposite order. With coloured stimulus patterns there was a possibility of impairment of judgement of TF by after-images and so one eye was used for inspection and the other for testing. One half the subjects observing in a particular order inspected with the left eye and tested with the right. The remainder used the opposite eyes for inspecting and testing.

An experimental session consisted of (*a*) a foretest using the TF, (*b*) an inspection period of 2 min. using the IF under the HD or BD condition, (*c*) the test using TF, (*d*) a rest period of 3 min. to allow recovery from the first inspection, (*e*) a second foretest, and (*f*) an inspection period of 2 min. using the second brightness difference condition. During the foretests, the subject was required to judge and report on the alignment of the two squares comprising the TF. Recording of responses was in terms of the shift of the lower square of the TF after inspection relative to its reported position during the foretest.

A shift of the lower square to the right indicated an FAE. All observations were made in daylight with the stimulus patterns illuminated from windows behind and to the left of the subject.

Preceding the foretest, the following instructions were given. 'Look carefully at the small brown dot with your left (right) eye. Now, without taking your eye from the dot, tell me about the two small squares to the right of it. Is the lower square in alignment with the one above it, further to the left or further to the right? It is best to judge the alignment of the squares from the two left-hand edges.' Before the inspection period, the following instructions were given. 'Now I want you to look carefully at the small dot for two minutes with your left (right) eye. Try to keep your eye steadily on the dot for the two minutes. Cover your other eye lightly with your hand. When the two minutes are up, I am going to show you the two small squares again. When I do so, I want you to change eyes, covering your right (left) eye and looking with your left (right) eye. Then tell me again how the bottom square appears relative to the top one. That is, is it in alignment, to the right or to the left? I want you to make this judgement immediately the squares are shown to you. Avoid hesitating and make your judgement as quickly as possible. It is important to report exactly what you see; how the squares actually appear to you, irrespective of how you think they should appear.'

Table 1. *Frequency of occurrence of figural after-effects, no after-effects, and effects opposite to the figural after-effect for the two conditions of brightness difference in the three experiments.*

(The judgements recorded represent changes following inspection.)

	Effect opposite to figural after-effect	No after-effect	Figural after-effect	Total
	Exp. 1			
Brightness difference	1	9	26	36
No brightness difference (hue difference)	2	11	23	36
Total	3	20	49	72
	Exp. 2			
Brightness difference	0	5	15	20
No brightness difference (hue difference)	0	5	15	20
Total	0	10	30	40
	Exp. 3			
Brightness difference	0	0	12	12
No brightness difference (hue difference)	0	0	12	12
Total	0	0	24	24

Results. Following inspection of the IF, the three possible judgements of the position of the lower TF square relative to its position during the foretest were either 'right', 'left' or 'no change'. The distribution of responses is shown in Table 1. Most of the displacements are to the right indicating a preponderance of FAE's over left displacements and no changes. For the two brightness difference conditions of the IF, the distribution of judgements over the three categories represents a significant departure from an equal number of judgements in each category. The χ^2 for this is significant at less than the 0·01 confidence level.

In order to compare the frequency of occurrence of responses under the HD condition with those using the BD inspection pattern, the responses in the opposite direction to those indicating an FAE and the 'no change' responses have been combined and compared with the FAE responses. The test developed by Stuart (1957) for the comparison of matched frequencies has been used in examining the difference in frequencies for the two conditions. This test has shown that there is no significant difference between the frequencies of FAE and no FAE responses for the BD and HD conditions. Thus, it can be concluded that, in this instance, the frequency of occurrence of the FAE with nearly equal brightness of IF and ground does not differ significantly from the frequency of occurrence when there is a marked brightness difference.

During the foretest preceding the HD inspection, 7 subjects judged the lower square of the TF to be displaced to the right of the square above it. These squares were objectively in alignment. During the foretest preceding the BD condition, 6 subjects made this judgement. Since an FAE was indicated when the lower TF square was displaced to the right after inspection of the IF, such foretest judgements pose a problem. In each of these instances, the subject was asked to compare the magnitude of the foretest displacements with that following inspection. If the subject indicated that the right displacement following inspection was clearly greater than that during the foretest, an FAE was indicated. If there was some uncertainty concerning the relative magnitudes of the two displacements, a 'no change' judgement was recorded.

Although subjects were not questioned concerning the relative magnitudes of displacement following inspection of the HD and BD patterns, 11 subjects who obtained an FAE under both conditions reported spontaneously that the displacement following the BD condition was greater than that for the HD pattern. One subject reported the reverse.

III. EXPERIMENT 2

Köhler's (1951) experiment dealing with brightness differences in the IF and the FAE was conducted as a demonstration during an experimental course. No details of this experiment are now available. The experiment to be described was carried out at the suggestion of Dr Köhler (private communication to the writer, 1957) using IF's and a TF (Fig. 1*b*) similar to those used in Köhler's original investigation.

Inspection and test patterns. Drawing paper was painted to match 2·5R 3/6 and 2·5BG 3/2. The Munsell value for the dark red was 3·47, and for the dark green 3·44. Thus, the difference in brightness between these two hues was 0·03 in Munsell terms. The IF's and fixation points were cut from the dark green paper. For the HD pattern, the IF and fixation point were mounted on the dark red ground and, for the BD pattern, the ground was white. The dimensions of the grounds for the two IF's and the TF were the same as in Exp. 1. In this instance, the visual angle subtended by the distance between adjacent IF and TF contours varied since the IF lines were slanted in relation to the lines of the TF (see Fig. 1*b*).

Subjects. There were 20 subjects, 15 males and 5 females. Ages extended from 17 to 42 years. The group comprised University and non-University personnel.

Design and procedure. The design of this experiment, together with procedure and viewing conditions, was similar to that in Exp. 1. In this instance, an FAE was indicated when, after inspection, the two parallel lines of the TF were seen as diverging at the top and/or converging at the bottom. Instructions to the subject before the foretest and inspection were similar to those in Exp. 1 but adapted for use with the different IF and TF.

Results. Results for the 20 subjects are shown in Table 1. Although 5 subjects under each condition of brightness difference failed to report an after-effect no subject obtained an after-effect opposite to that of the FAE. The distribution of judgements after inspection represents a significant departure from an equal number of judgements in the 'FAE' and 'no FAE' categories. Since the distributions of judgements are similar for both conditions, a test for the significance of differences between conditions has not been necessary. Generally, the outcome of this experiment again indicates that the FAE occurs with the same frequency with and without a difference in brightness in the inspection pattern.

IV. EXPERIMENT 3

The FAE's obtained with the IF's and TF's used in Exps. 1 and 2 could be due to a process of adaptation to a spatial norm of verticality in the manner suggested by Gibson (1933). It would be difficult to attribute an FAE involving an alteration in the size of a figure, especially that of a circle, to this adaptation process. In view of this, it was decided to conduct a third experiment using the 'size effect' rather than an FAE involving directional displacement.

Inspection and test patterns. The materials used in the construction of the patterns were sheets of commercially available blue and red cardboard which were found to have nearly equal luminous reflectances. In matching by eye to Munsell hues, it was found that the blue approximated 2·5 PB 5/8 and the red 5·0R 5/12. The true Munsell value for the blue was 4·52, and for the red 4·53. Both colours had a slightly glossy surface. The IF's and fixation points were cut from the blue. For the HD pattern, the ground was red and for the BD condition it was white. Since it was difficult to cut two annuli exactly equal in size these were drawn in black on a white ground. The IF and TF drawn to full scale and, in their positions relative to the fixation point and to one another, are shown in Fig. 1*c*. The visual angle subtended by the distance between adjacent contours of the TF and IF annuli was approximately 6 minutes.

Subjects. There were 12 subjects aged between 21 and 35 years. There were 9 males and 3 females all of whom were University staff or students.

Design and procedure. The design, procedure and viewing conditions were the same as for Exps. 1 and 2. Instructions before foretest and inspection were adapted for the IF's and TF used. An FAE was indicated when the right annulus of the TF appeared larger and/or the left appeared smaller.

Results. Results for the 12 subjects are shown in Table 1. In all cases, the right annulus of the TF was seen as larger than the left under both conditions of brightness difference in the IF. During the foretest preceding HD inspection, 3 subjects saw the right annulus of the TF as slightly larger than the left annulus. Preceding the BD condition, 4 subjects made this judgement. In each case, these subjects were requested to compare the relative difference in size of the two annuli before and after inspection. In all cases, the size difference following inspection was reported as markedly larger than the difference before inspection. Since the figures used in this experiment appeared to give a marked FAE, the 'right larger' foretest judgements did not constitute a problem in deciding whether or not an FAE had occurred. Because of the unequivocality of these results statistical treatment has been unnecessary.

V. DISCUSSION

Results from the three experiments amply confirm Köhler's (1951) contention that the FAE occurs when there is little or no difference in brightness between IF and ground. Although complete equality of the two parts of the inspection pattern was not achieved, the difference in all experiments, and especially in Exps. 2 and 3, was very small. The data show also that the frequency of occurrence of the FAE with nearly equal brightness of IF and ground does not differ significantly from the frequency obtained with a brightness difference. The possibility that these results apply only to certain classes of IF is not supported by the evidence, since the three IF's used varied in both form and colour. Whereas the stimulus patterns used in Exps. 1 and 2 would favour an adaptive process of the type suggested by Gibson (1933), those used in Exp. 3 would not. Irrespective of whether or not there are different types of visual after-effect, the present data demonstrate that, in instances where after-effects can only with difficulty be ascribed to spatial adaptation, the after-effect occurs in the absence of brightness differences.

It is of interest to consider briefly the failure of some subjects to report an FAE and the tendency of others to report an effect opposite to that of the FAE. George (1953) has shown that, in the case of FAE's involving changes in size, the occurrence or non-occurrence of an FAE depended largely upon the form of the IF. With an IF and TF similar to those used in Exp. 3, George demonstrated that the FAE occurred very much more frequently with an IF consisting of two annuli (one larger and one smaller than the TF annuli) than it did with an IF consisting of a single annulus. In view of this evidence, it is reasonable to assume that had another rectangle, placed diagonally opposite to the one used, been incorporated in the IF of Exp. 1 the number of recorded FAE's would have increased. Because of considerable individual differences in proneness to the FAE, some subjects either do not obtain an after-effect or obtain an effect opposite to the FAE when the stimulus patterns are such that they do not give rise to a strong or marked FAE.

The marked difference between the results reported by Hochberg & Triebel (1955) and the present findings deserve further consideration. Hochberg & Triebel failed to obtain an FAE when there were no differences in brightness within the inspection pattern. In Exp. 3, which is comparable in terms of the stimulus patterns used, an FAE was obtained under the HD condition with all subjects. The overall brightness of the stimulus patterns and the visual angles subtended by the figures were more or less similar in the two experiments. It is worth noting that, although the visual angles subtended by the distance between the adjacent contours of the IF and TF differ in the two investigations (that in Hochberg & Triebel's study being about one half that in Exp. 3), this factor would hardly affect the outcome. Data on the distance paradox reported by Köhler & Wallach (1944) and by Pollack (1958) show that the magnitude of the FAE is reduced by approximately 15% when the distance between adjacent contours of IF and TF is reduced by half. Only when this distance is reduced by more than half of the distances used here does the FAE begin to decrease to a marked degree. In any case, Hochberg & Triebel obtained an FAE in all cases with a brightness difference in the IF. In comparing the two experiments, the different colours in the stimulus patterns might be considered as a possible source of the inconsistent results. Throughout the three experiments reported here, however, these colours have varied widely, indicating that this factor is not crucial for the production of the FAE. The one marked difference between the two experiments lies in Hochberg & Triebel's use of an achromatic (grey) background for the IF, whereas in the present Exp. 3 the background was red. The results reported here are based on a larger number of subjects than that used by Hochberg & Triebel and argue strongly against their general conclusion that 'chromatic differences can neither modify nor generate the phenomenal after-effect'.

In formulating a neurophysiologically based theory to explain the FAE, Köhler & Wallach (1944) developed the notion of satiation, that is, an electronic state, which is said to arise through differential stimulation of adjacent cortical regions. Such differential stimulation at the cortex was held to be due to the greater frequency of impulses reaching one region than reaching the region immediately adjacent to it. Köhler & Wallach (1944) specifically excluded from their considerations instances in which areas differed in hue rather than in brightness (p. 328). At the time when the satiation hypothesis was developed, such a consideration would have been an embarrassment since little evidence was available to support a view that different hues give rise to different impulse frequencies in the optic tract. Recent investigations by Chang (1951) strongly support the view that there are three types of optic fibre, each of a different size and conducting velocity. Donner (1950) has also adduced evidence in favour of the view that the neural correlate of colour vision is a difference in the timing of impulses in the primary projecting system of the eye. If these data are fully substantiated, then the theories advanced so far to account for the FAE would not meet with undue difficulty in explaining the results reported here.

The general conclusion made possible by the present data is that the FAE, rather than being dependent on a brightness gradient in the inspection pattern, is dependent on the presence of a visible contour. It seems to be of little account whether this contour derives from a brightness gradient or a difference in hue. Evidence supporting such a view has been reported by Pollack (1958), Fujiwara & Obonai (1953) and Yoshida (1955). These investigators have shown that, as the difference in brightness between the inspected figure and its ground decreases, so the magnitude of the FAE is decreased. This suggests that, as the contours become less distinct, so the FAE grows weaker.

Finally, it is of interest to note that the present data, together with results reported earlier (Day, 1957, 1958), indicate that both the figural after-effect and the after-effects of seen movement occur when there is little or no difference in brightness within the inspection patterns.

The writer wishes to acknowledge the assistance given by Professor W. M. O'Neil in painting the papers to match Munsell hues, and by Mr W. Bleven of the National Standards Laboratory, Commonwealth Scientific and Industrial Research Organization, in measuring and computing the luminous reflectances of the hues.

REFERENCES

CHANG, H.-T. (1951). Functional organization of central visual pathways. *Res. Publ. Ass. Nerv. Ment. Dis.* **30**, 430–53.

CHANG, H.-T. (1956). Fiber groups in primary optic pathway of cat. *J. Neurophysiol.* **29**, 224–31.

DAY, R. H. (1957). The relation of brightness to the after-effect of seen movement. *Austr. J. Psychol.* **9**, 123–9.

DAY, R. H. (1958). The effect of small brightness differences on the after-effect of seen movement. *Quart. J. Exp. Psychol.* **10**, 166–70.

DONNER, K. O. (1950). The spike frequencies of mammalian retinal elements as a function of wave length of light. *Acta physiol. Scand.* **21**, Suppl. 72, 1–59.

FOX, B. H. (1951). Figural after-effects: 'Satiation' and adaptation. *J. Exp. Psychol.* **42**, 317–26.

FUJIWARA, K. & OBONAI, T. (1953). Quantitative analysis of figural after-effects. II. Effects of inspection time and intensity of light stimulus upon the amount of figural after-effects. *Jap. J. Psychol.* **24**, 114–20.

GEORGE, F. H. (1953). On the figural after-effect. *Quart. J. Exp. Psychol.* **5**, 128–35.

GIBSON, J. J. (1933). Adaptation, after-effect, and contrast in the perception of curved lines. *J. Exp. Psychol.* **16**, 1–31.

HOCHBERG, J. E. & TRIEBEL, W. (1955). Figural after-effects with coloured stimuli. *Amer. J. Psychol.* **68**, 133–5.

KÖHLER, W. (1940). *Dynamics in Psychology*. New York: Liveright.

KÖHLER, W. (1951). Relational determination in perception. In L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior*. New York: Wiley.

KÖHLER, W. & WALLACH, H. (1944). Figural after-effects: an investigation of visual processes. *Proc. Amer. Phil. Soc.* **88**, 269–357.

MUNSELL COLOR COMPANY, INC. (1950). *Munsell Book of Color* (Opposite Hue Edition). Baltimore.

NICKERSON, D. (1948). Colour and its description. *Bull. Amer. Ceramic Soc.* **27**, 47–55.

POLLACK, R. H. (1958). Figural after-effects: quantitative studies of displacement. *Aust. J. Psychol.* **10**, 269–77.

STUART, A. (1957). The comparison of frequencies in matched samples. *Brit. J. Stat. Psychol.* **10**, 29–32.

YOSHIDA, T. (1955). An experimental study of figural after-effects. *Jap. J. Psychol.* **23**, 287–8.

(*Manuscript received* 8 *March* 1958)

INTER-SENSORY PERCEPTUAL SET

By GEORGE A. TALLAND

Harvard Medical School and Massachusetts General Hospital

A set to perceive words representing a given category, more specifically animal names, was induced by visual presentation of some other words drawn from the same class. Measures were taken to prevent vocalization of these visual stimuli. Under these conditions, the facilitatory effect of a set was found to be no less marked than it is when auditory material is used both for establishing and for testing the set. In either instance, the effect of set was demonstrated by comparison with correct report of the same words, when heard scattered among other words, or as a continuous list but without previous indication of their selective composition.

The finding is presented in support of a central theory of mental set, for neither events in the receptor or in the effector areas could account for it without some intervening process in the brain.

I. Introduction

One of the minor controversial topics in psychology has been about the location of mental sets. Does the neuro-physiological process which exercises a selective control in motor behaviour or in perception involve the effector or the receptor apparatus of the nervous system, or does it operate centrally? Mowrer (1940, 1941) presented the findings of his experiments on motor sets in support of the central theory, but his interpretation of the studies was subsequently disputed by Gibson (1941). Some recent research in the neuro-physiology of perception (Kuffler & Hunt, 1952; Galambos, 1955; Granit, 1955) holds out promise of resolving the dilemma, very probably by demonstrating that the mechanism operates at both levels. Meanwhile, psychologists, too, have studied more intensively the selective processes in perceptual expectancy, i.e. set in a sense distinct from readiness for a motor response. Their experiments usually require the subject to reproduce stimuli presented at or near the threshold level. A common method of establishing sets or expectancies is that of defining, for the subject, the class from which the forthcoming stimuli will be drawn, e.g. pictures of fruit or names of capital cities. Experiments have shown that identical figures can be perceived as a letter or a number (Bruner & Minturn, 1955), as an animal or a mountain scene according to the expectancy, and moreover may not be recognized as a figure perceived in accordance with the alternative set (Zangwill, 1937).

Most of these experiments have used tachistoscopic material, and even though the expectancy is established by auditory verbal instruction, confirmation proceeds in the same modality as its test, that is, visually. Presumably, the set operates at a neuro-physiological level which is common to both sense modalities, i.e. centrally. Nevertheless, it could be argued that, since all stimulation affects one special sense organ, the set operates by lowering excitatory thresholds in that receptor. Concepts, of course, must be formed by some central process, but once formed they could theoretically operate peripherally, as suggested by the work of Lacey & Smith (1954), whose subjects' conditioned anticipatory autonomic responses generalized by means of unverbalized conceptual associations.

The purpose of this experiment was to test the central theory of set, by comparing the facilitatory effect of visually and aurally induced expectancies on the accuracy of auditory perception. Inducing an expectancy means the implicit establishment of a set by examples or cues, as contrasted with explicit instruction or warning.

II. PROCEDURE

Forty-eight staff members of the Department of Psychiatry of the Massachusetts General Hospital served as subjects. Twelve each were arbitrarily assigned to the experimental group and to three control groups. All groups were tested on auditory material, the experimental group having been presented first with visual stimuli suitable to induce a set congruent with the test material. Control group VA (for visual-auditory) was also first exposed to visual stimuli; however, these were not likely to induce a set. Control group AS (for auditory set) was given first the same set inducing words as the experimental group, but by means of auditory presentation. Control group AN (auditory: no set) was also presented with auditory material only, but in a manner unlikely to induce a set.

The test material consisted of words drawn from one conceptual class, animal names. The effect of a set was determined by correct reproduction of these words presented on a recorded tape, at or near threshold level. Subjects sat at a distance of 4–5 ft. from a Revere recorder, those in the experimental group and in control group VA in front of a Gerbrands tachistoscope. They were instructed to 'look into' the tachistoscope throughout the entire experiment, even though they were presented with the visual stimuli only over a relatively short phase of it. Auditory stimulation continued from the beginning to the end of the experiment, interrupted for a break after each word so that subjects could reproduce it without interfering with the next stimulus word. Their instruction was to report immediately the word they had heard. No indication was given about the type or class from which the words would be drawn, or that they would belong to any one category. Subjects were also warned that the words spoken on the recorder would be only just loud enough to be intelligible, and that therefore they might not clearly understand every one. They were asked, nevertheless, to repeat without delay and in a clear voice whatever they had heard or thought they had heard.

For the experimental group and control group VA, the procedure consisted of three phases. The first phase served to calibrate the instrument to each subject's individual threshold of intelligibility. This was done by presenting him with a *pilot list* of 36 mixed words, and by adjusting the volume and tone controls to positions at which the subject reported correctly as many words as he failed to.

In the second phase, both groups were presented with a further 12 mixed words on the tape recorder. After each of the last six items of this *pre-list*, subjects were also shown another word in the tachistoscope. The instruction preceding this phase was the same as that given earlier, i.e. to report each word they heard on the tape just as it sounded to them. This time, however, subjects were also warned that they would see words in print in the tachistoscope, and told not to read these aloud. The words were printed in heavy type of about $\frac{1}{8}$ in., and displayed in a white field 1 in. high and 3 in. wide, in the centre of a black panel which, in the Gerbrands tachistoscope, is about 22 in. from the subject's eye. Tachistoscopic exposure was synchronized with the auditory stimulus, and lasted 0·25 sec. This was a shorter time span than that required for pronouncing the dissyllabic spoken word, but above the threshold of accurate visual perception.

The experimental group was shown 6 animal names in the tachistoscope; the control group VA was shown 6 mixed words, each different from those presented aurally. Neither mixed list included animal names. The 12 words on the tape were: nothing, island, complain, himself, report, common, almost, impact, open, suppose, human, between. In conjunction with the last six, the control group read the following: money, coarse, bolt, bat, rule, gear. The matched animal list shown to the experimental group listed: monkey, horse, wolf, rat, mule, bear.

The purpose of this *pre-list* was to establish, if possible, a set for auditory perception, merely by visually presenting names drawn from one class of words, and preventing the subject from vocalizing these by requiring him to pronounce a different word while reading each of these names. Subjects invariably observed the instruction to report only the aurally presented stimuli.

The third phase of the experiment tested the establishment and operation of a mental set, by presenting on the recorder for vocal report a *test list* of 18 animal names: elephant, antelope, deer, beaver, chipmunk, pony, racoon, sheep, tiger, skunk, fox, leopard, zebra, reindeer, gorilla, rabbit, camel, pig.

Control groups AS and AN received the same instructions as the experimental group, except that they did not have to look into the tachistoscope. They started the experiment with the standard *pilot list* of 36 mixed words. Control group AS was given the same 6 animal names for a *pre-list* as the experimental group, but presented on the tape; these were followed by the first 12 items of the *test list*. Control group AN was presented with these 12 animal names interspersed among words drawn from other categories or from a mixed selection, evenly spaced in a total list of 108. All tapes were recorded by the same voice, and the 12 words common to the *test list* of the experimental group and control groups AS and VA were re-recorded from a master tape. The effect of a set was tested by the number of correct reports, each subject's total on the *test series* being treated as a score. Records were also made of incorrect responses congruent with the set in operation, i.e. animal names not actually presented.

III. RESULTS

The experimental group and control group VA gave precisely the same correct report from the *pre-list*, 7·83; variances of the two distributions of scores were homogeneous. The mean correct report from the first 12 animal names of the *test list* was 9·50 in the experimental group, 6·92 in control group VA, 9·00 in control group AS, and 6·67 in control group AN. Having satisfied the criterion of homogeneity of variances, t ratios were calculated, and these were found to be significant at the 1% level between the experimental group and control groups VA and AN ($t = 4{\cdot}20$ and $6{\cdot}56$), and also between the latter two and control group AS ($t = 4{\cdot}14$ and $4{\cdot}47$). The difference between the means of the experimental group and control group AS ($t = 0{\cdot}83$), and between control groups VA and AN ($t = 0{\cdot}63$), were not statistically significant. The mean correct report of items 13–18 of the *test list* was 4·50 in the experimental group, and 4·33 in control group VA; the difference between these is not significant.

In the first 12 items of the *test series* the experimental group reported a total of seven animal names other than those spoken on the tape; the corresponding totals were six in control group AS, and three in control group VA. Owing to the small frequencies, statistical testing was inapplicable; however, the trend is evidently parallel to that indicated by the effect of expectancy on accuracy of report.

IV. DISCUSSION

The experiment has demonstrated that a perceptual set or expectancy for a certain class of objects can be induced by means of another sense modality than the one involved in perception. To this extent, the present experiment merely tested, under different conditions, a principle inherent in a number of tachistoscopic studies, in which a set for visual stimuli was established by auditory verbal instruction. This experiment has extended the principle in two respects: first, by establishing the expectancy implicitly rather than by explicit instruction; secondly, by providing conditions which aimed at preventing the subject from translating the set into the sense modality in which its effect was tested. In previous tachistoscopic experiments there was no means of checking whether the subject produced visual images of objects or symbols congruent with the set while receiving this by auditory verbal instruction. In the present study, his performance interfered with the vocalization of examples of the class of words he was to be tested on.

A previous study, Talland (1958) tested the effect of a set by more accurate auditory perception than occurs without such a set and demonstrated that this effect is significant whether the set had been established 'implicitly' or 'explicitly' by auditory means. The present experiment shows that, using the same test based on auditory stimulation, the perceptual set can be induced visually, and that its effect is no less marked than that of sets similarly induced through the auditory presentation of cues. A control situation in which the visual stimuli provide no indication of the class from which the test material will be drawn, makes for no more accurate perception of these test words than does their isolated presentation amidst other words which do not belong to their class. In both these latter situations accurate perception occurred significantly less often than after establishment of a set. The question, whether verbal report is a satisfactory operational measure of perception, can only be answered with the observation that in every situation subjects tended to respond with an incorrect word rather than not to respond at all.

The finding that a visually induced 'implicit set' is as effective in facilitating auditory perception as an aurally induced set, is put forward in support of a central theory of the locus of mental sets. In establishing an expectancy of animal names for subjects of the experimental group, nothing had occurred outside a subject's organism that could have lowered thresholds in or otherwise affected the sense organ which subsequently became involved in the test, i.e. the ear. The same argument applies to the effector as well as to the receptor organ, for subjects had no opportunity to vocalize the set-inducting stimuli. Expectancy, therefore, could not have been aroused as a backlash (Freeman, 1948) or otherwise derivative process of the motor response, an explanation favoured by peripheralists. Whatever the neural mechanism of set, by the evidence of this experiment it would appear to operate at a level which is common to two, or more, different sense receptor systems, i.e. centrally.

REFERENCES

BRUNER, J. S. & MINTURN, A. L. (1955). Perceptual identification and perceptual organization. *J. gen. Psychol.* **53**, 21–8.

FREEMAN, G. L. (1948). *The Energetics of Human Behaviour*. Ithaca: Cornell University Press.

GALAMBOS, R. (1955). Suppression of auditory nerve activity by stimulation of efferent fibers to the cochlea. *Fed. Proc.* **14**, 53.

GIBSON, J. J. (1941). A critical review of the concept of set in contemporary experimental psychology. *Psychol. Bull.* **38**, 781–817.

GRANIT, R. (1955). Centrifugal and antidromic effects on ganglion cells of the retina. *J. Neurophysiol.* **18**, 388–411.

KUFFLER, S. W. & HUNT, C.C. (1952). The mammalian small nerve fibers: A system for efferent nervous regulation of muscle spindle discharge. *Proc. Ass. Res. Nerv. Dis.* **30**, 214–47.

LACEY, J. I. & SMITH, R. I. (1954). Conditioning and generalization of unconscious anxiety. *Science*, **120**, 1043–52.

MOWRER, O. H., RAYMAN, N. M. & BLISS, E. L. (1940). Preparatory set (expectancy)—and experimental demonstration of its 'central' locus. *J. Exp. Psychol.* **26**, 357–72.

MOWRER, O. H. (1941). Preparatory set (expectancy)—further evidence of its 'central' locus. *J. Exp. Psychol.* **28**, 116–33.

TALLAND, G. A. (1958). The effect of set on accuracy of auditory perception. *Brit. J. Psychol.* **49**, 117–30.

ZANGWILL, O. L. (1937). A study of the significance of attitude in recognition. *Brit. J. Psychol.* **28**, 12–17.

(*Manuscript received* 6 *May* 1958)

OBJECTIVE-TEST ASSESSMENT OF THE PRIMARY PERSONALITY DIMENSIONS IN MIDDLE CHILDHOOD*

By RAYMOND B. CATTELL
University of Illinois
AND RICHARD W. COAN
University of Arizona

Factor analysis was applied to the scores of 164 children in the 6–8-year age range on 111 objective-test variables. A simple-structure solution was obtained for 16 factors, and factor scores were correlated with 49 additional variables. The inclusion of marker variables patterned after variables in earlier batteries permits a comparison of present factors with factors found in a number of studies for older children and adults. Every factor appears to correspond to a dimension known from earlier work. Eleven or twelve factors can be identified with considerable confidence. In the remaining factors, the correspondence to the expected loading pattern is somewhat less satisfactory.

I. Aim of the investigation

Over the past several years, statistical analysis of experimental data has been in progress at the University of Illinois to define personality structure at the adult and pre-adolescent levels (Cattell, 1957). In more recent research, the analysis has been carried to the 6–8-year and 4–5-year levels. At these ages, as in earlier work, the approach has been threefold. Factors have been independently isolated in the media of behaviour ratings (Cattell & Coan, 1957*a*, *b*), questionnaire data (Cattell & Coan, 1955; Coan & Cattell, 1958, 1959) and objective-test measurement.

Results for the rating and questionnaire realms have been reported elsewhere (Cattell & Coan, 1957*a*, *b*, 1958; Coan & Cattell, 1958, 1959). Our present concern is with objective test findings at the 6–8-year level. The subject population in this phase of the study was the same as that defined for the behaviour ratings (Cattell & Coan, 1957*a*, *b*). Thus, there were four first-grade classes and four second-grade classes with a total potential N of 198. Statistical analysis, however, was confined to 164 subjects for whom we managed to secure a complete set of test data.

In age, these children ranged from 6 years 3 months to 8 years 11 months, with a mean of 7 years 5 months. There were 82 girls and 82 boys.

II. The test battery

A deliberate attempt was made to include in the present study *marker variables for every factor previously found in the objective-test realm for pre-adolescents or adults*, additionally to the creation of new tests. There were, however, practically no tests in the batteries for older groups which were suitable for use, without suitable modification, at the 6–8-year level. In subsequent matching, either with adults or with Peterson's 4–5-year group (Cattell & Peterson, 1959) allowance must be made appropriately for certain differences of the marker variables from their analogues in earlier batteries. The new tests were suggested both by various theoretical considerations and by direct observations of children's behaviour in miniature situations.

---

* The research reported in this article was supported by a grant from the Department of Public Welfare of the State of Illinois.

Efforts were made, wherever possible, to construct tests in a form suitable for group administration. In all, 43 groups tests were devised. These utilized a variety of situations, employing verbal, visual, auditory and social stimuli. Tests requiring the use of apparatus or pure motor and verbal responses were designed for individual administration. There were 16 individual tests altogether.

The group tests, together with the questionnaire described elsewhere (Cattell and Coan, 1958; Coan & Cattell, 1958, 1959) were distributed over ten testing sessions. The individual tests were distributed over two sessions. Every testing session lasted about 45 minutes. In group testing, each class was tested as a unit, and the tests were administered, as far as possible, in a fixed sequence on ten consecutive school days. The order of the two individual sessions necessarily varied somewhat, and various children were given individual tests before and after the group tests.

Since every child included in our objective-test sample received a 200-item questionnaire, it proved convenient to include in our score matrix a few objectively definable response patterns (notably deviant-response tendency, position-response tendency, intra-factor consistency, and consistency with correlation) assessed from questionnaire responses. The questionnaire items as such have been analysed separately (Cattell & Coan, 1958; Coan & Cattell, 1958, 1959). A recent, as yet unpublished research shows that these factors have continuity with those of the adult 16 Personality Factor Questionnaire. Their relation to the objective test factors in the present study has not yet been determined, though work has been in progress for a year estimating the factor scores on a common group in order to determine these relations. For present purposes the test battery yielded 155 scores. To these were added five variables (numbered 156 to 160 below) based on information external to our testing.

In the following list, there is a capsule description of each variable. In addition to the variable number in our present matrices, we have indicated a Master Index (M.I.) number to facilitate cross-reference with earlier studies and the Test Encyclopedia kept at the Laboratory of Personality Assessment. For purposes of economy, the description of the variable will be omitted when the present test and its administration manifest no significant operational departure from its predecessor and the reader is then referred to the O–A Battery (Cattell *et al.* 1955) where a description and specimen are available. When possible, we have also included an indication of the reliability of the variable. In every case, the measure presented is a split-half consistency coefficient corrected by the Spearman–Brown formula.

The following variables were employed in this study:

Matrix number	M.I. number	Variable
1	360	*Consistency of values: in hexads.* Rank consistency within sets of value elements, each set of four elements being subjected to all six possible paired choices and constituting one scorable item. $R = 0\cdot28$
2	113	*Acceptance of reality principle* (G9, cf. 3). $R = 0\cdot28$
3	246	*Respect for authority* (G34, G9). $R = 0\cdot37$
4	219	*Honesty in admitting common frailties* (G41). $R = 0\cdot79$
5	100	*Pessimism over doing good* (G16). $R = 0\cdot39$
6	112	*Ratio favourable/unfavourable self-reference* (G24). $R = 0\cdot25$
7	321	*Restrained reading preferences* (G25). $R = 0\cdot78$

Matrix number	M.I. number	Variable
8	21	*Questionable story preferences* (G5). $R = 0.84$
9	108	*Confident assumption of skill in untried performance* (G22). $R = 0.67$
10	116	*Critical severity vs. indulgent standards* (G26). $R = 0.77$
11	29	*Carping criticism and mischievous humour* (G40). $R = 0.92$
12	103	*Amount considered possible for self in given time* (G19). $R = 0.70$
13	211	*Susceptibility to annoyance* (G38). $R = 0.63$
14	211*a*	*Susceptibility to annoyances: ego threats.* Based on responses to a series of verbal items like those for M.I. 211, but involving content more directly and specifically relating to situations threatening the perceived adequacy of the subject. $R = 0.75$
15	67	*Extremity of viewpoint* (G9, G10). $R = 0.76$
16	361	*Confident assumption of knowledge.* Readiness to assert possession of knowledge that subject is very unlikely to possess. $R = 0.78$
17	363	*Story completion: ego control vs. malevolent external intervention.* In response to incomplete stories, tendency to choose alternative in which principal character masters the given problem in preference to an alternative in which the principal character becomes the unwilling victim of another (physical or animate) agent. $R = 0.50$
18	364	*Story completion: internal vs. external control.* Like M.I. 363 (number 17 above), except that alternative completion may involve either a benevolent or a malevolent agent. $R = 0.35$
19	368	*Dominant story identification.* Expressed preference for being a child who manifests superiority or competitive skill as against a child who is said to achieve satisfaction of some other kind. $R = 0.48$
20	39	*Colour/form ratio in sorting* (G(A) 1). $R = 0.83$
21	227	*Picture preferences: colour/form ratio* (G(A) 4). $R = 0.31$
22	370	*Visual memory.* Correct identification within a set of pictures of those which have been previously presented. $R = 0.73$
23	371	*Pictures interpretation: ego control vs. benevolent external intervention.* Tendency to choose as a solution to a pictured problem situation a solution in which the central character achieves mastery through his own efforts, as against a solution in which the problem is overcome by some external agent. $R = 0.35$
24	372	*Picture interpretation: ego control vs. malevolent external intervention.* Like number 23, except that the alternative solution is one in which the central character becomes the victim of a malevolent external agent. $R = 0.31$
25	374	*Preference for disturbing sounds.* Expressed preference for sounds which are startling, eerie, or suggestive of distress or destruction, as against more neutral sounds. $R = 0.82$
26	167	*Efficiency of immediate memory* (G9). $R = 0.37$
27	31	*Fluctuation of attitudes* (G9, G10). $R = 0.39$
28	120	*Ratio accuracy/accomplishment at maximum speed* (G48). $R = 0.74$
29	375	*Circle cancellation performance.* Overall speed and accuracy in a task requiring subject to mark the circles, and only the circles, in a set of varied figures. $R = 0.45$
30	309	*Speed of line length judgements* (G45). $R = 0.91$
31	377	*Numerousness judgement accuracy.* Speed and accuracy in deciding in two-choice items which member contains more figures. $R = 0.48$
32	378	*Quantity judgement accuracy.* Speed and accuracy in deciding which of each pair of geometrical figures is larger. $R = 0.75$
33	203	*Self-confidence in quantitative skill* (G35). $R = 0.83$
34	379	*Goodness of maze performance* (I). Speed and accuracy of pencil-maze performance. $R = 0.41$
35	380	*Carelessness of maze performance* (I). Extent of gaps and excursions in pencil-maze performance, relative to total distance covered. $R = 0.46$
36	206	*Perceptual analysis: Gottschaldt* (G37). $R = 0.79$
37	379	*Goodness of maze performance* (II). Like number 34 above, but based on a different maze test. $R = 0.95$
38	380	*Carelessness of maze performance* (II). Like number 35 above, but derived from a different maze test. $R = 0.89$
39	101*b*	*Excess aspiration level over performance* (I) (G17). Measured in a pencil-maze task. $R = 0.65$
40	145	*Responsiveness aspiration level to performance* (I) (G17). Measured in a pencil-maze task. $R = 0.62$
41	381	*Design preferences: Complex vs. simple.* Expressed preference, in response to paired abstract line drawings, for designs of greater linear complexity. $R = 0.91$
42	382	*Speed of evaluative judgements.* Number of items marked in given time in design-preference task
43	143	*CMS total score* (G14). Adapated as a group paper-and-pencil task, with a complexity of marking instructions suited for 6-year-olds. $R = 0.91$

Matrix number	M.I. number	Variable
44	266	*CMS ratio of final (fast)/initial (slow) total scores* (G14). $R = 0{\cdot}95$
45	133	*Criticalness of judgement* (G8). $R = 0{\cdot}99$
46	383	*Inhibition of criticism.* Extent to which expressed criticism is reduced by the anticipation that subject will have to present oral justification before his class of his critical marks. $R = 0{\cdot}71$
47	384	*Picture preference: strange vs. familiar.* Expressed preference, in response to paired pictures of common objects, for strange version of the object rather than a familiar version. $R = 0{\cdot}85$
48	385	*Skill in rapid shape comparison.* Accuracy and speed in a task requiring decisions as to whether paired shapes are the same or different. $R = 0{\cdot}87$
49	247	*Expansiveness of copied design* (G(A) 5). $R = 0{\cdot}74$
50	386	*Expansiveness of figure drawing.* Area covered by drawing of a person. $R = 0{\cdot}61$
51	387	*Form level of figure drawing.* Goodenough score for a full-figure drawing of a man
52	388	*Ratio expansiveness other/self figure drawing.* Area utilized when child is asked to draw a picture of 'a person' relative to area utilized when he is asked to draw himself
53	389	*Whole/part ratio in figure drawing.* Tendency to respond with a whole-person drawing, rather than merely a head, when asked to draw a person. $R = 0{\cdot}79$
54	9	*Ideomotor tempo* (G50). Unspeeded production of X's for a limited period of time
55	391	*Untidiness of colouring.* Extent to which the child, in colouring a set of printed pictures, tends to overshoot the boundaries of the figures. $R = 0{\cdot}62$
56	392	*Ratio complete vs. partial colouring.* A ratio of number of pictures completely coloured to number partially coloured in free colouring task. $R = 0{\cdot}26$
57	393	*Original drawings.* Tendency to produce and colour original drawings, when given this alternative, instead of colouring pictures which have already been drawn
58	400	*Preference for dangerous vs. safe situations.* Tendency in a pencil maze with alternative paths to select paths representing dangerous situations or activities. $R = 0{\cdot}80$
59	401	*Preference for competitive vs. non-competitive situation.* Tendency in a pencil maze with alternative paths to select paths representing competitive situations. $R = 0{\cdot}35$
60	404	*Level of attention.* Accuracy of marking in accordance with a set of easy but very monotonous oral directions. $R = 0{\cdot}61$
61	405	*Loss of attention.* Same task as for number 60, scored as a ratio of later to earlier performance during the session. $R = 0{\cdot}64$
62	406	*Readiness to volunteer comments.* Number of general-discussion questions to which the child responds in classroom situation by raising his hand to indicate desire to express an opinion
63	271	*Verbal fluency* (G43). $R = 0{\cdot}68$
64	273	*Proportion of fluency on self* (G43). $R = 0{\cdot}83$
65	316	*Relative fluency on own characteristics* (G13). $R = 0{\cdot}40$
66	30	*Ratio of self-criticism to criticism of others* (G13). $R = 0{\cdot}95$
67	407	*Slow simple decision time.* Time taken to give oral response to two-choice preference questions. $R = 0{\cdot}58$
68	409	*Time ratio for few/many alternatives.* Oral response time for two-choice preference items relative to time for four-choice items. $R = 0{\cdot}53$
69	270	*Arm-shoulder swing tempo* (I17). Number of circles executed in four 20-second trials. $R = 0{\cdot}98$
70	269	*Leg-circling tempo* (I16). Time taken for ten circles with each foot. $R = 0{\cdot}98$
71	412	*Motor restraint.* Time taken to trace a set of four lines when instructed to perform the task as slowly as possible without actually stopping. $R = 0{\cdot}97$
72	413	*Ratio final/initial motor restraint.* For same task as in number 71, ratio of time for last two trials to time for first two trials. $R = 0{\cdot}58$
73	414	*Generalized muscular tension.* Pneumographically recorded amplitude of pressure on a rubber bulb held in non-writing hand during motor restraint task. $R = 0{\cdot}73$
74	175	*Reaction time with irregular warnings* (I1). $R = 0{\cdot}85$
75	40	*Reactions to false RT signals* (I1). $R = 0{\cdot}63$
76	71	*Complex RT: errors of response* (I1). $R = 0{\cdot}55$
77	72	*Ratio of simple/complex RT* (I1). $R = 0{\cdot}40$
78	263	*Reaction time: scatter* (I1). $R = 0{\cdot}32$
79	286	*Ratio visual/auditory RT.* $R = 0{\cdot}32$
80	340	*Formboard performance time.* Time taken to complete a formboard with simple geometrical inserts. $R = 0{\cdot}85$
81	339	*Mirror drawing time.* Time taken to draw a line between two given positions with hand visible only via the mirror. $R = 0{\cdot}48$

Matrix number	M.I. number	Variable
82	357	*Mirror drawing distance.* Total distance (length of line) covered in attempt to perform mirror-drawing task. $R = 0{\cdot}43$
83	419	*Involvement in task.* Reaction of apparent surprise or disappointment, as against indifference or indeterminate emotional response, recorded when a tower that child is constructing of blocks topples. $R = 0{\cdot}54$
84	420	*Single tower construction time.* Time taken to build a tower of a single column of blocks, from onset to point at which tower collapses. $R = 0{\cdot}33$
85	421	*Cautious anticipation.* Discrepancy between actual toppling height of block tower built by child and height at which child indicates that it is not 'safe' to build higher. $R = 0{\cdot}78$
86	422	*Height of single-tower construction.* Height attained in construction of a single-column tower of blocks. $R = 0{\cdot}57$
87	101*b*	*Aspiration level relative to performance* (II) (G17). Measured in a situation requiring construction of the highest possible single-column block tower. $R = 0{\cdot}50$
88	145	*Responsiveness of aspiration level to performance* (II) (G17). Measured in tower-construction task. $R = 0{\cdot}73$
89	423	*Planfulness.* Degree of planfulness, according to an eightfold criterion, manifested in child's attempt to construct, with blocks, a bridge resembling a schematic drawing. $R = 0{\cdot}66$
90	424	*Bridge construction time.* Time taken to construct, with blocks, a bridge resembling a schematic drawing shown to the child. $R = 0{\cdot}45$
91	282	*Unstructured drawings: total fluency of perception* (G20). $R = 0{\cdot}89$
92	105	*Unstructured drawings: perception of threatening objects* (G20). $R = 0{\cdot}65$
93	425	*Unstructured drawings: perception of aggressive implements.* Number of aggressive implements seen. $R = 0{\cdot}50$
94	6	*Slow ideomotor speed* (II) (G49). Time taken in a colour naming task. $R = 0{\cdot}73$
95	2	*Motor perceptual rigidity: reverse colour naming.* Speed ratio of time taken to name alternate colour, in response to given colour, to straight colour naming. $R = 0{\cdot}56$
96	2	*Motor-perceptual rigidity: reverse colour naming: errors.* Number of incorrect responses in reverse colour naming task. $R = 0{\cdot}44$
97	8	*Alternating perspective* (G3). $R = 0{\cdot}93$
98	171	*Imaginative perception of hidden objects* (G33). $R = 0{\cdot}79$
99	7	*Slowness of Gestalt closure* (G2). $R = 0{\cdot}47$
100	146	*Accuracy of Gestalt closure* (G2). $R = 0{\cdot}47$
101	426	*Distractibility in picture task.* Proportion of time spent looking elsewhere when presented with paired pictures to examine. $R = 0{\cdot}80$
102	427	*Attention to aggressive vs. non-aggressive pictures.* Time spent looking at pictures portraying aggressive behaviour paired with non-aggressive pictures in picture inspection task. $R = 0{\cdot}52$
103	428	*Attention to disturbing vs. non-disturbing pictures.* Time spent looking at pictures portraying emotionally disturbing situations paired with non-disturbing situations in picture-inspection task. $R = 0{\cdot}50$
104	429	*Preference for aggressive vs. non-aggressive pictures.* Number of aggressive pictures selected as preferred to their non-aggressive counterparts in picture-inspection task. $R = 0{\cdot}77$
105	430	*Preference for disturbing vs. non-disturbing pictures.* Number of disturbing pictures selected as preferred to their non-disturbing counterparts in picture inspection task. $R = 0{\cdot}73$
106	431	*Recall for aggressive vs. non-aggressive pictures.* Number of aggressive pictures relative to number of non-aggressive pictures recalled immediately after picture inspection task. $R = 0{\cdot}81$
107	435	*Reminiscence vs. rapid forgetting.* Total number of pictures recalled after 15-minute delay relative to number immediately recalled following picture-inspection task. $R = 0{\cdot}36$
108	152	*Tendency to agree* (G9). Total number of 'yes' responses over a variety of 'yes...no' verbal tests. $R = 0{\cdot}68$
109	436	*Position response tendency.* Degree of departure from random sequence in responses to dichotomous questionnaire items presumed to be so arranged that correlations between adjacent items are negligible. $R = 0{\cdot}46$
110	437	*Irrelevant marking.* Number of pages in group questionnnaire and verbal tests containing marks (doodling) not governed by test instructions. $R = 0{\cdot}84$
111	439	*Deviant responses.* Number of questionnaire items on which child gave a response given by fewer than 20% of the total sample. $R = 0{\cdot}66$
112	65	*Logical consistency of attitudes* (G9). Each scoring unit consisted of four spaced two-choice items constituting a syllogistic form for which the overall pattern of responses can be scored as consistent or inconsistent. $R = 0{\cdot}06$

Matrix number	M.I. number	Variable
113	362	*Story completion: ego control vs. benevolent external intervention.* In response to incomplete stories, the tendency to choose alternative in which the principal character masters the given problem in preference to an alternative in which the problem is solved by an external agent. $R = 0{\cdot}18$
114	365	*Tendency to choose morally positive definitions.* When presented with equally correct definitions of a given word, tendency to choose an alternative connoting moral virtue in preference to a morally neutral alternative. $R = 0{\cdot}17$
115	366	*Tendency to choose morally negative definitions.* Tendency to choose a definition connoting moral defect in preference to a morally neutral alternative. $R = 0{\cdot}17$
116	367	*Tendency to choose moral definitions.* Tendency to choose definitions with a moral tone, positive or negative, as against neutral alternatives. $R = 0{\cdot}53$
117	369	*Memory for person vs. object pictures.* Correct identification of previously seen pictures of people relative to correct identification of previously seen pictures of objects. $R = 0{\cdot}10$
118	373	*Picture interpretation: internal vs. external control.* In response to a pictured problem situation, tendency to choose an interpretation or outcome in which the central character masters the situation, as against an interpretation or outcome in which an external agent, benevolent or malevolent, is in control. $R = 0{\cdot}04$
119	38	*Excess consonant over dissonant recall* (G9). $R = 0{\cdot}10$
120	34	*Immaturity of opinion* (G9, G10). $R = 0{\cdot}10$
121	193, 194	*Shift via prestige suggestion.* Tendency to alter questionnaire responses on a re-examination in a direction toward what are said to be the responses of high prestige groups and away from responses said to be those of low prestige groups. $R = 0{\cdot}10$
122	13	*Oscillation* (G4). Derived from circle-cancellation task
123	312	*Accuracy of line judgements* (G45). $R = 0{\cdot}89$
124	376	*Ratio correct/attempted line judgements.* Ratio of total items marked according to instructions to total number of items attempted in task requiring line-length comparisons. $R = 0{\cdot}98$
125	237	*Ratio difficult/easy* (G42). $R = 0{\cdot}13$
126	51	*Index of carefulness* (G(A) 2). Derived from a task involving differential marking of assorted geometric figures. $R = 0{\cdot}18$
127	307*a*	*Speed in shape comparison* (G44). $R = 0{\cdot}77$
128	6	*Ideomotor speed* (I) (G49). Production of X's in speeded performance
129	390	*Number colours used in colouring task.* Total number used in a task involving free colouring of a set of printed line pictures
130	394	*Total drawings coloured.* Number pictures attempted in colouring task
131	395	*Adherence to sequence.* Ratio of number pictures that child attempted to colour to item number of last attempted picture
132	396	*Ratio colouring familiar/strange pictures.* Ratio of number pictures of people or things in familiar form coloured to number in strange form coloured
133	397	*Ratio colouring easy/difficult pictures.* Ratio of number of pictures with relatively large component areas coloured to number of finely articulated pictures coloured
134	398	*Ratio colouring people/objects.* Ratio of number of pictures of people used in colouring task to number of pictures of things
135	399	*Preference for company of parent vs. peer.* Tendency in a pencil maze with alternative paths to choose paths representing activities done with parents rather than with peers. $R = 0{\cdot}00$
136	402	*Preference for solitary vs. social situations.* Tendency in a pencil maze with alternative paths to select paths representing solitary pursuits. $R = 0{\cdot}23$
137	403	*Preference for company of mother vs. father.* Ratio of number of parent choices in pencil maze involving company of mother to number of parent choices involving company of father
138	109	*Ratio happy/unhappy associations* (G23). $R = 0{\cdot}18$
139	408	*Decision time ratio for simple/complex motives.* Ratio of time spent responding to two-choice items involving simple appetitive goals to time spent responding to similar items relating to more complex, socialized motives. $R = 0{\cdot}06$
140	410	*Ranking consistency for simple motives.* Ranking consistency of four simple-motive goal elements, judged from responses to six two-choice items which contain all possible paired comparisons
141	411	*Ranking consistency for complex motives.* Like number 140, but involving goals of more complex, socialized motives
142	28	*Endurance of difficulty* (G51). Derived from a task requiring the child to hold each leg in turn in a horizontal position. $R = 0{\cdot}35$
143	45	*Ratio final/initial output* (G51). Time ratio for later to earlier trials in leg-holding task. $R = 0{\cdot}04$
144	415	*Ratio initial/final muscular tension.* Increase in muscular tension, measured as in number 73, during successive trials in motor restraint task. $R = 0{\cdot}02$

Matrix number	M.I. number	Variable
145	5	*Ratio irregularly/regularly warned RT* (I1). $R = 0{\cdot}06$
146	165	*Approval gain and disapproval loss* (I18). Derived from formboard performance. $R = 0{\cdot}23$
147	416	*Caution: initial building on broad base.* When asked to build a high tower of blocks, tendency to begin building on a base of more than one block
148	417	*Counting of blocks.* Counting of blocks while building, when asked to build block tower as high as possible
149	418	*Acceptance of block task.* Expressed interest in repeating block-building task, when child was asked after each trial whether he would like to try again
150	170	*Perception of hidden pictures: number correctly seen.* (G33). $R = 0{\cdot}76$
151	432	*Recall for disturbing vs. calm pictures.* Ratio of number disturbing pictures recalled in picture-inspection task to number of their non-disturbing counterparts recalled. $R = 0{\cdot}15$
152	433	*Reminiscence for aggressive pictures.* Ratio of number agressive pictures recalled after 15-minute delay following picture-inspection task to number immediately recalled. $R = 0{\cdot}31$
153	434	*Reminiscence for disturbing pictures.* Like number 152, but based on disturbing pictures. $R = 0{\cdot}11$
154	438	*Intra-factor consistency.* Extent of agreement between two sets of questionnaire factor-score estimates for given child
155	440	*Consistency with common personality structure.* Extent to which direction of responses within highly correlating pairs of questionnaire items is consistent with direction of correlation for the total sample. $R = 0{\cdot}71$
156	64	*Sex*
157	277	*Age*
158	517	*Grade level in school*
159	441	*Absences. Number of absences for the school year*
160	442	*Metropolitan reading readiness score.* Since the children in the present sample were generally given this test at the beginning of the school year in the first grade, this provides a measure of relative aptitude within the child's class, rather than relative aptitude within the total sample.

III. TREATMENT OF THE DATA

Before statistical analysis was undertaken, the scores for each variable were converted to normalized T-scores (except in the case of the few dichotomous variables). It was not possible to apply factor analysis to the complete set of variables, since available computer techniques set a limit of 111 variables for a single analysis. Accordingly, 111 variables were selected on the basis of three criteria: (*a*) reliability, (*b*) relative experimental independence from other variables to be included, and (*c*) relevance to and suitability for marking various hypothesized personality factors.

The variables thus selected have been assigned matrix numbers from 1 to 111 in the above list. The correlation matrix among them was obtained in terms of Pearson product-moment coefficients. For the dichotomous variables, of course, these reduce to phi and point-biserial coefficients. The resultant matrix was factor analysed by the complete centroid method.

On the basis of various criteria of completeness of factor extraction, it was originally decided that 15 factors should be retained for rotation. A sixteenth factor was inserted, however, when an additional unmistakable hyperplane appeared in the course of rotation. The final oblique solution, obtained after numerous rotations beyond the oblimax analytical solution, showed, by Bargmann's (1954) test, significance at the P 0·001 level, only in one case (factor 10) descending to the P 0·01 level. All factors are, of course, oblique, since this alone permits a thorough pursuit of the simple structure. However, the correlations among the reference vectors in the C_R matrix are all very slight, only 4% of them being above 0·3, which is roughly what might be expected by chance from a matrix of this size.

From the final solution, a complete set of regression weights was derived, and a set of factor-score estimates was obtained for every subject. These estimates were then correlated with the remaining variables (variables 112–160 above) to provide information regarding their factor composition.

IV. FINDINGS

We shall now examine the factors one by one, taking them in order of diminishing contribution to the variance of the totality of *directly factored variables*. In what follows we shall present for each factor a list of its salient variables in order of loading.* This list of more highly loaded variables has normally been cut off at a lower limit of 0·25, though loadings down to the hyperplane are presumably significant, but of little value for estimating the factor. In addition, *all* the variables required to have significant loadings if the match hypothesized is correct are included with the given salients (in italics) *whatever* their actual loadings may be. Finally each table ends with a separate section for the correlations of the estimated factor with the extension variables (112–160). Both because some factors, where indicated, have been slightly rotated after the estimation, and also because of attenuation, these values would be likely to be higher if the variables had been included directly in the factorization. For each variable, the description has been worded to agree with the sign of the loading and with the positive pole of the factor.

Table 1. *Factor* 1. *Assertiveness;* U.I. 16

Matrix number	Variables	Loading
34	Rapid and accurate maze performance	+71
30	*Fast speed of line judgements*	+68
37	Rapid and accurate maze performance	+64
42	Rapid evaluative judgements	+46
29	Rapid and accurate circle cancellation	+45
32	Rapid and accurate quantity judgement	+44
80	Rapid formboard performance	−40
35	Careless ('essentials only') maze performance	+37
38	Careless ('essentials only') maze performance	+37
36	Rapid and accurate perceptual analysis	+35
54	Fast ideomotor tempo	+35
43	High CMS total score	+33
69	*Fast tempo of arm-shoulder swing*	+22
91	*Many objects perceived in unstructured picture*	+08
125	*High ratio distance of difficult to easy mazes*	+05
95	Low motor rigidity	−02
120	*Maturity of opinion*	−02
	Extension variables	
123	*High accuracy of line-length judgements*	+86
127	*Rapid shape comparison*	+73
158	Second grade in school, rather than first grade	+61
150	Many hidden objects correctly seen	+55
128	Fast ideomotor speed	+52
157	Older, rather than younger	+42
126	High index of carefulness	+37

Factor 1, the hyperplane of which is significant beyond the 0·001 level (cf. Bargmann, 1954), has good loadings on most markers for U.I. 16 (2), no trace of the required pattern for which occurs in other factors. (All have correct sign, but rigidity and maturity are so modified from the adult form that their loadings are negligible.) The confidence of the match is enhanced by the factor of largest variance here coinciding with that of largest variance in earlier studies (Cattell, 1957)—hence number U.I. 16 therein—and by the general character of the new variables continuing the quality of rapidity and accuracy ('Competent assertiveness' or 'Desire to achieve') found before. The link pointed out earlier (Cattell, 1957) with one of Thurstone's perception speed factors is strongly confirmed. This assertiveness factor increases rapidly with age.

---

* The centroid matrix, transformation matrix, rotated reference-vector matrix, and a matrix of correlation between factor-score estimates and the extension variables have been deposited with the American Documentation Institute. Order Document No. 5979 from American Documentation Institute, 1719 N St. N.W., Washington 6, D.C., remitting $1.25 for 35 mm. microfilm or $1.25 for 6 by 8 in. photocopies.

Table 2. *Factor 2. Timidity inhibition;* U.I. 17

Matrix number	Variables	Loading
104	Low preference for aggressive pictures	−72
105	Low preference for disturbing pictures	−65
8	*Absence of questionable reading preferences*	−62
7	*Many restrained preferences*	+54
25	Low preference for disturbing sounds	−53
58	Low preference for dangerous situations	−52
11	*Little mischievous humour*	−51
102	Relatively little attention given to aggressive pictures	−50
103	Relatively little attention given to disturbing pictures	−40
42	Slow evaluative judgements	−28
106	Relatively poor recall for aggressive pictures	−28
9	Little confident assumption of skill	−26
10	Low critical severity	−26
28	*High accuracy relative to amount accomplished at maximum speed*	+11
77	*Much slowing of RT by complex instructions*	−11
92	*Low perception of threatening objects in unstructured drawings (A reversed marker)*	−07
145	*High ratio of regular to irregular warned reaction time*	+02
	Extension variables	
156	Female, rather than male	−48
155	High consistency with common personality structure	+31
134	Low tendency to colour pictures of people rather than pictures of objects	−24
136	Low preference for social, rather than solitary situations	+20

It is interesting that, again, the factor of second highest variance here is that found second in previous evaluations (Cattell, 1957, p. 239); for this contains, without rival, the markers for U.I. 17, though in two or three cases the analogue test marks inadequately. Again the new variables confirm the original hypothesis of 'inhibition' or 'timidity', e.g. the highest loadings are in avoidance of aggressive and disturbing pictures and sounds, and of dangerous situations. Agreement with U.I. 17 in the 4–5 study (Cattell & Paterson, 1959) is only moderate. The substantial correlation (not identity) with female sex is noteworthy, as also the interest in inanimate things relative to people.

Although factor 3 is the unquestioned match in this series for U.I. 21, Exuberance (Cattell, 1957), having all markers correct in sign and three well loaded, there is some shift in emphasis from adults. There is greater prominence of aggression, plus fear of aggression, and lesser 'insuggestibility to authority'. The 'general clinical' resemblance of U.I. 21 to U.I. 16—along with the quite significant difference when markers are scrutinized—noted in previous researches, persists here. The factor level has some tendency to increase in second grade.

Factor 4 alone has the necessary markers for U.I. 26, with which the new variables continue to agree, depicting self control, ego awareness and ego involvement against an asocial, autistic 'fiddling'. However 'dynamic momentum' is poorly loaded, agreeing with

Table 3. *Factor* 3. *Exuberance;* U.I. 21

Matrix number	Variables	Loading
91	High fluency for perceived objects in unstructured drawings	+73
98	High imaginative perception of hidden objects	+58
99	High speed of Gestalt closure	−55
93	High perception of aggressive implements in unstructured drawings	+52
92	High perception of threatening objects in unstructured drawings	+48
63	*High verbal fluency*	+32
62	High readiness to volunteer comments	+29
81	Rapid mirror drawings	−26
108	*Tendency to indicate agreement rather than disagreement*	+11
30	*Fast speed of line judgement*	−09
120	*Immaturity of opinion*	+12
142	*High dynamic momentum*	−06
	Extension variables	
150	Many hidden objects correctly seen	+60
123	High accuracy of line judgements	+36
128	*Fast ideometer speed*	+33
158	Second grade, rather than first grade	+33
127	*Rapid shape comparisons*	+32
130	Many drawings coloured in given time	+28
157	Older, rather than younger	+28
126	High level of carefulness	+26

Table 4. *Factor* 4. *Self sentiment control;* U.I. 26

Matrix number	Variables	Loading
75	Few reactions to false RT signals	−42
76	*Few erroneous reactions in complex RT*'s	−35
66	Little criticism of others relative to self-criticism	+30
64	*High proportion of fluency on self*	+28
83	Much apparent emotional involvement in block task	+28
108	High tendency to agree	+26
63	*Low total fluency on topics*	−20
95	*Low motor-perceptual rigidity (speed)*	−12
92	*High ratio of threatening objects seen in unstructured drawings*	+05
44	*Good maintenance of CMS performance under more speeded conditions*	+19
20	*Preference for form over colour in sorting*	−08
	Extension variables	
158	Second grade, rather than first grade in school	+34
123	High accuracy of line-length judgements	+31
118	Predominance of internal control over external control in picture interpretations	+23
133	Low tendency to colour easy, rather than difficult, pictures	−21

Table 5. *Factor* 5. *Corticalertia;* U.I. 22

Matrix number	Variables	Loading
74	*Fast RT with irregular warnings*	−68
75	*Many reactions to false RT signal*	+57
76	Many errors of response in RT task with complex instructions	+53
77	Much slowing of RT with complex instructions	−38
78	*Low variance in reaction time*	−31
35	Much carelessness of maze performance	+30
15	High extremity of viewpoint	+30
59	High preference for competitive situations	+29
91	*Many objects perceived in unstructured drawing*	+12
97	*High-speed alternating perspective*	+14
	Extension variables	
145	*Low ratio of irregular to regular warned RT*	−10
122	*Much oscillation of performance*	+19
142	*Low dynamic momentum*	−12
156	Male, rather than female (inversion)	+15
143	*High ratio final to initial output*	+02
118	*Predominance of internal control over external control in picture interpretations*	+20

our two earlier studies rather than Brogden (see Cattell, 1957). Higher school grade is confirmed, as in our studies and Brogden's.

In factor 5, all nine markers for U.I. 22 have correct signs, though number 143 is negligible. It is noteworthy that when markers are absolutely unchanged, as here in the reaction time performances, they lose no loading and even enhance a little. The essential 'corticalertia' quality of sheer high speed, with some carelessness and easy triggering is retained in the new variables, with a suggestion also of keyed-up competitiveness. However, the former suggestion in student groups of female sex association is here, in the general population, unsustained. (Presumably selection of women college students.) With U.I. 22 in the Peterson 4–5-year old study there is additional agreement on 76, 'Many errors in maintaining complex reaction time sets', 'rapid form board performance' and 'high fluency on self'.

Table 6. *Factor* 6. *Anxiety versus assured character;* U.I. 24

Matrix number	Variables	Loading
13	*High annoyability—all situations*	+52
14	*High susceptibility to annoyance involving ego threats*	+50
45	*High criticalness of judgement*	+46
3	*Low respect for authority*	−36
69	Fast arm-shoulder swing tempo	+35
46	Much inhibition of criticism under social pressure	+33
17	Predominance of malevolent external intervention over ego control in story completion	−26
4	*Willingness to admit common frailties*	+03
108	*Tendency to agree*	+01
8	*Absence of questionable reading preferences*	−14
9	*Little confident assumption of skill*	−09
70	*Fast leg-circling tempo*	−18

Table 7. *Factor* 7. *Dourness;* U.I. 33

Matrix number	Variables	Loading
9	Little confident assumption of skill	−42
12	*Little considered possible for self in given time*	−41
55	Little untidiness in colouring	−38
110	Little irrelevant marking on answer sheets	−38
13	High susceptibility to annoyance	+34
6	*Predominance of unfavourable over favourable self-reference*	−34
2	Much acceptance of reality principle	+33
59	Low preference for competitive situations	−32
47	Low preference for strange versions of common objects	−31
24	Predominance of malevolent external intervention over ego control in picture interpretations	−30
58	Low preference for dangerous situations	−29
39	Little excess of aspiration level over performance	−28
87	Little excess of aspiration level over performance	−27
85	High cautious anticipation	+21
44	*High ratio of final to initial performance*	+07
95	*Low motor-perceptual rigidity (speed)*	−03
	Extension variables	
155	High consistency with common personality structure	+34
156	Female, rather than male	−26
160	*High MRR score*	+24
158	Second grade in school, rather than first grade	+20

Factor 6 has the markers for the well-known anxiety factor (Cattell, 1957) though tendency to agree, and willingness to admit frailties, drop sharply from their position in the adult factor, just as they do in the 4–5-year old anxiety pattern (Cattell &

Peterson, 1959). However, irritability, lack of confidence and criticalness remain central. The increased tempo, in 69 and 70, is probably tied to the individual testing situation.

Factor 7 has the markers for U.I. 33 (three good, two poor: no other factor has the right loading sign pattern) and the general meaning is overwhelmingly that previously recognized in the 'Dourness' factor. The new variables bring self criticism (9), self restraint (55, 110), irritability (13), suspicion of external malevolence (24), dislike of the strange (47), the dangerous (58), the competitive (59), caution (85), realism (2)—and the essence of the Scottish 'dour temperament' concept. As before (Cattell, 1957) there is some positive association with verbal-numerical ability (achievement component?), and, in our present sample, a slight association with femininity. The converse pole is obviously one of confident, expansive, immature optimism and lack of restraint.

Table 8. *Factor* 8. *Exvia versus Invia;* U.I. 32

Matrix number	Variables	Loading
86	Low tower construction	−52
89	Low planfulness	−46
84	Short time in tower construction	−42
4	Little admission of common frailties	−34
110	Little irrelevant marking on answer sheets	−29
96	Many errors in reverse colour naming	+26
44	Low maintenance of CMS performance under more speeded conditions	−24
65	*High fluency on own characteristics*	+13
91	*Much fluency of perception on unstructured drawings*	+12
9	*Much self confidence on untried performance*	+10
100	*Low accuracy on Gestalt perceptual completion*	−04
1	Low consistency in Hexads	−22
	Extension variables	
113	Predominance of benevolent external intervention over ego control in picture interpretations	−29
150	Few hidden objects correctly seen	−21
156	Female, rather than male	−20

In factor 8, the markers for the extroversion-introversion factor, U.I. 32 (now called Exvia-Invia, in the factor realm, to dissociate it from unverified and unspecifiable clinical concepts in extraversion-intraversion), have, as usual, slight loadings similarly (but negative) on the U.I. 24 Anxiety factor, and, positively, on the U.I. 21 Exuberance factor. The present factor is the only other one with correct sign loadings for all variables, but they are fainter than would be desired for a confident match. However, the meaning of the new non-marker variables: impulsive, uncritical versus planful, restrained, internally-consistent; fits the Exvia-versus-Invia concept very well (Cattell, 1957). The consistency of values test (1) has its highest loading here, in the introvert direction.

Factor 9 alone has the asthenia factor markers with nine out of ten correct in sign, though they have dropped, in all but four, within the hyperplane. Presumably this peculiar combination of asthenia with aggression and authority problems is one changing a good deal through childhood. It has always been a factor of broad influence and low saturation (Cattell, 1957); but the asthenia continues to show, e.g. in the slow mirror drawing, low speed of decision in visual comparisons (also found in Peterson's U.I. 28 (Cattell & Peterson, 1959)); the aggression and discouragement in the 'criticalness combined with carelessness', and the sociable compliance in other tests.

Factor 10 has only one sign reversal in ten markers for U.I. 19 and the matching loadings are very good. The 4–5-year old study (Cattell & Peterson, 1959) agrees on the newly loaded variables 'few errors in false RT' and 'high variance RT'. Its whole sense is exactly that of the practical precision, control and critical sense stated in the compulsive superego concept of U.I. 19. As usual in U.I. 19, there stands at the opposite pole a quality of distractibility, poor organization and inability to keep firm mental sets. Curiously the hyperplane has the lowest significance of any; $P < 0{\cdot}01$, but $> 0{\cdot}001$.

Table 9. *Factor* 9. *Asthenia;* U.I. 28

Matrix number	Variables	Loading
81	Slow mirror drawing	+46
82	Much meandering in mirror drawing	+45
28	High accuracy relative to amount accomplished at maximum speed	+39
29	High speed and accuracy in circle cancellation	+33
65	*High fluency on own characteristics*	+32
62	High readiness to make voluntary comments	+28
10	High critical severity	+27
89	Low planfulness	−27
110	Much irrelevant marking on answer sheets	+20
108	*Tendency to agree*	+01
13	*High over-all annoyability*	+04
15	*Much extremity of viewpoint*	+13
99	*Fast speed of Gestalt closure*	−04
3	*Much authority submission*	+04
121	*Ego weakness, much shift to successfuls*	+06
119	*Low ratio of consonant/dissonant recall*	−07
	Extension variables	
158	First grade in school, rather than second grade	−39
126	Low level of carefulness	−35
123	Low accuracy of line-length judgements	−34
157	Younger, rather than older	−32
128	Slow ideomotor speed	−30
118	Predominance of external control over internal control in picture interpretations	−28
150	*Few hidden picture correctly seen*	−27
127	Slow shape comparison	−23
122	Low variance of circle-cancellation performance	−22
112	*Low logical consistency of attitudes*	−20

Table 10. *Factor* 10. *Critical practicality;* U.I. 19

Matrix number	Variables	Loading
95	*Low motor-perceptual rigidity*	−56
103	Relatively little attention to disturbing pictures	−42
101	Little distractibility in picture-inspection task	−40
15	Low extremity of viewpoint	−35
36	*Good analysis performance on Gottschaldt figures*	+34
94	Slow ideomotor speed	+33
45	*High criticalness of judgement*	+32
46	Much inhibition of criticism under social pressure	+28
102	Relatively little attention to agressive pictures	−28
75	*Few reactions to false RT signals*	−27
76	*Few errors in maintaining complex reaction-time sets*	−04
20	*Low tendency to sort by color rather than form (inverted)*	−14
10	*High critical hostility of judgement*	+14
100	*High percentage accuracy in perceptual Gestalt completion*	+13
126	*High index of carefulness*	+17
	Extension variables	
160	High MRR score	+29
155	High consistency with common personality structure	+27
150	*Many hidden objects correctly seen*	+26
156	Female, rather than male	−24
113	Predominance of ego control over benevolent external intervation in story completions	+21
146	Low tendency to improve performance with approval and to get worse with disapproval	−21

The few markers available for U.I. 31 have non-hyperplane loadings in factor 11, but one is sign-inverted, so the match is tentative. However, as in the Wary Realism concept (Cattell, 1957) the new variables continue to show some insecurity together with a strong bid for internal control, and we are inclined to accept the identification as a basis for further research.

Table 11. *Factor* 11. *Possibly, wary realism;* U.I. 31

Matrix number	Variables	Loading
18	Predominance of internal control over external control in story completions	+61
17	Predominance of ego control over malevolent external intervention in story completions	+60
105	High preference for non-disturbing over disturbing pictures	−38
16	Little confident assumption of knowledge	−36
23	Predominance of ego control over benevolent external intervention in picture interpretations	+32
7	Much restrained preferences	+28
47	High preference for familiar, over strange versions of common objects	−27
78	High variance in reaction time	+27
15	*Little extremity of viewpoint*	−11
33	*Low self-confidence in computational skill*	−12
12	*Much considered possible in given time for self*	−10
27	*Much fluctuation of attitudes (inversion)*	+20
	Extension variables	
118	Predominance of internal control over external control in picture interpretations	+38
156	Male, rather than female	+24
132	High tendency to colour strange, rather than familiar, pictures	−21

Table 12. *Factor* 12. *Hypomanic smartness;* U.I. 18

Matrix number	Variables	Loading
66	*Much self-criticism relative to criticism of others*	+52
48	*Rapid and accurate shape comparison decisions*	+43
65	Relatively much fluency on own characteristics	+36
21	High preference for form over colour in pictures	−35
103	Much attention to non-disturbing, relative to disturbing pictures	−34
88	Little decrement of aspiration level in response to performance	+29
64	Relatively high fluency on self	+27
43	*High total score on CMS*	+19
44	*Low ratio of initial to final performance on CMS*	+05
20	*High ratio colour to form in sorting*	+06
29	*Rapid and accurate circle cancellation*	+07
31	*Rapid and accurate number judgement*	+22
32	*Rapid and accurate quantity judgement*	+12
34	*Rapid and accurate maze performance*	+04
46	*Little inhibition of criticism under social pressure*	−07
119	*Low ratio of consonant/dissonant recall*	−09
55	*Little untidiness in colouring*	−09
121	*Ego weakness: shift toward successfuls*	−10
	Extension variables	
156	Male, rather than female	+24
123	High accuracy of line-length judgements	+22
158	Second grade in school, rather than first grade	+21

Factor 12 has thirteen markers all correct in sign and the match is also, in all new test meanings, excellent. The fact that eight markers nevertheless have subsignificant loadings suggests, with other similar experience, that more weight should be given to sign pattern (probability of match here equals 1 in 2^{12}) than loading magnitudes. A secondary resemblance to U.I. 26 may be noted, but is not enough to occasion doubt. Avoidance of colour, and disturbance, in pictures is consistent with the ambitious self-direction, seen also in correlation of this factor with questionnaire factor N (Cattell, 1957). Self concern and irrepressible assertion continue in the new variables found in the factor at this age.

The match of factor 13 with U.I. 25 cannot be made with confidence, for the Gestalt performance loadings are negligible, yet no other factor shows the marker sign pattern. Without prejudice to adopting or rejecting Eysenck's view of U.I. 25 (−) as psychoticism, we can see, at the hypothesized 'Reality Contact' pole, more adjustment to realities of space, time and complexity here, and lower 'imaginative' (closure) speed, but the contrast of realism and imagination stresses the imaginative speed variables less than with older subjects. With Peterson's 4–5-year-old U.I. 25 there is agreement (lower loadings than

Table 13. *Factor* 13. *Possibly, realism versus psychotic tendency;* U.I. 25

Matrix number	Variables	Loading
53	Little tendency to draw head rather than whole person	+40
78	Small variance in reaction time	−38
50	*Small figure drawing*	−35
77	Much slowing of RT with complex instructions	−34
74	Fast reaction time with irregular warnings	−33
49	*Small size of drawing (low expansiveness) copying design*	−32
44	Little decrement of CMS performance under more speeded conditions	+31
12	Little considered possible for self in given time	−27
97	*Slow alternating perspective*	−18
94	*Fast ideomotor speed*	−16
99	*Slow speed of perceptual Gestalt closure*	+02
100	*Low accuracy of Gestalt closure*	+00

Table 14. *Factor* 14. *Comention;* U.I. 20

Matrix number	Variables	Loading
79	Auditory RT fast relative to visual RT	+35
57	Much tendency to make original drawings rather than colour presented drawings	+31
56	Predominance of partially coloured over completely coloured pictures in free colouring task	−29
86	High tower construction	+29
97	Fast alternating perspective	+28
62	High readiness to volunteer comments	+26
103	Much attention to non-disturbing, relative to disturbing, pictures	−26
73	Much generalized muscle tension	+21
15	*High extremity of viewpoint*	+20
108	*Tendency to indicate agreement rather than disagreement*	+17
45	*High criticalness of judgement*	+12
85	*Much cautious anticipation*	+17
102	*Much attention to non-aggressive, relative to aggressive, pictures*	−18
8	*Absence of questionable preferences in reading*	−03
4	*Willingness to admit common frailties*	+03
2	*Much acceptance of reality principle*	+10
3	*Much authority submission*	+07
	Extension variables	
150	Few hidden objects correctly seen	−28
159	Many absences from school	+21
114	*Tendency to chose an alternative connoting moral virtue in preference to a morally neutral alternative*	+10
120	*Immaturity of opinion*	−08
119	*High ratio of consonant/dissonant recall*	+07

salients, at the positive pole, on 'low fluency on self', 'high general fluency', and 'low extremity of viewpoint').

All markers for U.I. 20 fit in factor 14 only, being all correct in sign, and, for most, clear of the hyperplane. However, they have all dropped relatively to the adult pattern and new well-loading variables have taken over. However, it must be confessed that the new salients do not make the factor more comprehensible, which suggests they must be confirmed before the difficult required hypothetical constructs can be discussed. A tendency spontaneously to busy oneself may explain 57, 56 and 62, as well as the highest loading found anywhere in 'muscle tension', 73, which fits the notion of 'drive to activity'. The notion of inner activity matches the temperamental, high genetic determination found, and suggests that the cultural and moral positiveness is a secondary association, as the very low loadings here indicate.

Table 15. *Factor* 15. *General intelligence;* U.I. 1

Matrix number	Variables	Loading
35	Careful maze performance	−65
38	Careful maze performance	−60
43	High CMS total score	+35
22	Good visual memory	+33
94	High ideomotor speed	−31
77	Little lengthening of reaction time with complex instructions	+30
111	Few deviant questionnaire responses	−30
26	Good memory for statements	+29
29	Rapid and accurate circle cancellation	+29
60	High level of attention	+29
63	High verbal fluency	+29
71	Much motor restraint	+29
36	Good perceptual analysis	+28
	Extension variables	
160	High MRR score	+48
127	Quick shape comparisons	+47
123	High accuracy of line judgements	+42
150	Many hidden objects correctly seen	+42
126	High level of carefulness	+41
128	Fast ideomotor speed	+41
158	Second grade in school, rather than first grade	+40
125	Poor performance on difficult, relative to easy, mazes	−38
155	High consistency with common correlation trend	+29
157	*Older, rather than younger*	+26

No markers were placed in the personality study for the general ability factor, *g*; but as usual it makes its appearance, in factor 15. It may be considered marked by good memory, judgement, older and second-grade designation, ideomotor speed, total C.M.S. and the nearest thing to an intelligence test in the study, 160, with its highest loading here. Only the high loadings on 35 and 38 are a little odd. A remote second to general intelligence would be U.I. 23, Neural Reserves, most markers for which have some loading. But co-operativeness of intelligence and U.I. 23 has been noted before (Cattell, 1957) and may be of theoretical importance for the concept of neuroticism (cf. Cattell, 1957).

The greatest total appropriate loading for U.I. 23 lies in factor 16. Most signs are correct, as shown, though the loadings are sometimes negligible. Peterson's study (Cattell & Peterson, 1959) similarly indicates that in small children the *non*-neurotic makes *more errors* in the rapid colour-naming rigidity test. U.I. 23 seems in general, as here, to be a factor of quite small variance in normal populations (Cattell, 1957). The new salients, which at the neurotic pole are: many deviant responses (111), abnormally increased decision time with many alternatives (68), drawing the self large relative to others (52), and increased decision true for complex motives (139), fit the concept of ego weakness and may prove valuable new tests of the neuroticism factor.

Thus, we have found factors which correspond very clearly to known factors bearing the following universal index (U.I.) numbers: 1, 16, 17, 18, 19, 20, 21, 22, 24, 26 and 33. Factors which appear to correspond to U.I. 23, 25, 28, 31 and 32 do not reach an entirely satisfactory level of matching, though the sense of the new variables in these factors fully sustains matching made on the basis of low-loading marker variables. The fact that loadings for marker variables drop below the level generally found in cross-validation with groups of the same age points to the existence of systematic changes in personality factor patterns with age, such as occur, here and elsewhere, in the general intelligence factor. Incidentally, these results suggest that there is greater constancy in sign pattern than in loading, for in these cases of reduced loading the sign agreement is far beyond the usually accepted limit of chance ($P = 0{\cdot}05$). (Eight markers agreeing in sign with the hypothesized pattern, as typical here, gives $P = 0{\cdot}008$.)

Table 16. *Factor* 16. *Neuroticism* U.I. 23 (−) *versus neural reserves*

Matrix number	Variables	Loading
20	High tendency to sort by colour rather than form	+46
111	Many deviant questionnaire responses	+42
4	*High tendency to admit common frailities*	+40
52	Drawing of self large relative to drawing of other person	−39
67	Quick decision time	−31
8	Many questionable story preferences	+26
68	Long decision time for many alternatives relative to few alternatives	−23
22	*Poor visual memory*	−23
28	*Low accuracy relative to amount accomplished at maximum speed*	−21
26	*Poor memory for statements*	−19
14	*High susceptibility to annoyance involving ego threats*	+15
21	*Preference for form over colour in pictures*	+10
96	*High motor perceptual rigidity (many errors)*	+10
24	*Predominance of ego control over malevolent external intervention in picture interpretations*	+10
73	*Much generalized muscular tension*	+09 >
39	*Little excess of aspiration over achievement*	−03
95	*Low motor perceptual rigidity (speed) (inverted)*	+02
142	*Low endurance of difficulty*	−11
125	Good performance on difficult mazes relative to easy	+25
139	Long decision time for complex, relative to simple motives	−25
160	Low MRR score	−25

The general conclusion may be drawn that in objective tests the same personality dimensions as are known in adults persist down to at least the 7-year-old level. The identification of these dimensions has been accompanied by the invention of new tests suitable for 7-year-olds, and showing such validities as should permit practical batteries to be constructed for measuring source traits in the experimental study of personality in young children.

The writers wish to express their indebtedness to the school system of Decatur, Illinois, for generously giving the substantial school time and supervision necessary for advance in this area, and to the Mental Health Fund Committee, Dept. of Public Welfare, State of Illinois, for supporting the project.

REFERENCES

BARGMANN, R. (1954). Signifikanzuntersuchungen der einfachen Struktur in der Factoren-analyse. *Mitteilungsblatt für Mathematische Statistik*, Sonderdruck.

CATTELL, R. B. (1957). *Personality and Motivation Structure and Measurement.* New York: World Book.

CATTELL, R. B., BAGGALEY, A. R., CHECOV, L., COGAN, E. A., FLINT, D., GRUEN, W., HUSEK, E., MEELAND, T., SAUNDERS, D. R. & SCHIFF, H. (1955). *Handbook for the Objective-Analytic Personality Test Batteries.* Champaign, Illinois: Institute for Personality and Ability Testing.

CATTELL, R. B. & COAN, R. W. (1957*a*). Child personality structure as revealed in teachers' behavior ratings. *J. Clin. Psychol.* **13**, 315–27.

CATTELL, R. B. & COAN, R. W. (1957*b*). Personality factors in middle childhood as revealed in parents' ratings. *Child Develpm.* **28**, 439–58.

CATTELL, R. B. & COAN, R. W. (1958). Personality dimensions in the questionnaire responses of six and seven year olds. *Brit. J. Educ. Psychol.* **28**, 232–42.

CATTELL, R. B. & PETERSON, D. R. (1959). Personality structure in 4- and 5-year olds: in terms of objective tests. *J. Clin. Psychol.* (in the Press).

COAN, R. W. & CATTELL, R. B. (1958). Reproducible personality factors in middle childhood. *J. Clin. Psychol.* **14**, 339–45.

COAN, R. W. & CATTELL, R. B. (1959). The development of the Early School Personality Questionnaire. *J. Exp. Educ.* (in the Press).

(*MS. received* 21 *April* 1958)

POSITIVE AND NEGATIVE INFORMATION IN MATCHING PROBLEMS

By MARGARET DONALDSON

Psychology Department, Edinburgh University

Two tasks are used, for the successful performance of which it is necessary to prefer negative information to its formally equivalent positive counterpart. The subjects are 19 children aged 14 years to 14 years 6 months, who are observed individually, and 90 students in a Scottish teachers' training college who are tested as a group. Results confirm previous findings that formal equivalence of positive and negative information does not ensure psychological equivalence. Further, in certain circumstances, positive and negative information may be regarded as equivalent in certain respects but not in others. Although negative statements are recognized as combining to 'equal' a positive one, the equality may be incomplete since this derived positive is not always allowed the same value for purposes of further inference as a directly stated positive. Bruner, Goodnow & Austin (1956) suggest that avoidance of 'in-the-head' transformation may explain the tendency to shun negatives; but this does not seem adequate to account for the present findings. These authors also speak of 'inability or unwillingness' to use negative information. In this study, evidence both of inability and of unwillingness is obtained.

I. INTRODUCTION

In the study of problem solving it is possible, in certain circumstances, to calculate the formal value of a piece of information and compare this with the psychological value of the information for a given subject. It is also possible to equate, for formal value, information given positively and negatively, and then compare, for psychological value, these two different modes of communication. Much of the work on this topic has been done within the framework of the concept attainment problem.

Studies of the utilization of negative instances in concept learning have been made by Smoke (1933), Hovland & Weiss (1953) and Bruner *et al.* (1956). Smoke found no significant difference between the time required for learning from positive instances and that required for learning from a combination of positive and negative instances; but, as Hovland & Weiss point out, Smoke does not say what information was provided by his negative instances and so no measure of formal as opposed to psychological information value can be obtained from his data. This defect is remedied in the other two studies, both of which provide evidence that positive information is more efficiently utilized in concept attainment than is negative information.

Hovland & Weiss express the view that a major remaining research task is that of determining the factors responsible for this finding and suggest that perceptual factors must play some part since: 'Positive instances have the required characteristics directly perceptible. Negative instances, on the other hand, do not have the correct combination of characters in direct view.' This suggestion emphasizes an important feature of the concept attainment problem, namely, that the characters are normally presented to the subject in all possible combinations and the information consists in telling the subject whether each combination, as it appears in visible association before him, is a positive or a negative instance of the concept. But there is a situation which, from the point of view of logical structure, resembles the concept attainment one, and yet differs from it in this respect: that the characters are originally presented in complete dissociation from one another (if indeed they are presented non-verbally at all) and the information consists in instructions to the subject to link them or keep them apart in various ways. This occurs in any matching problem. Evidence of unequal difficulty in the handling of positive and negative information in matching problems will then be relevant to the Hovland & Weiss hypothesis concerning the role of perceptual association; and evidence is provided by a matching problem experiment of Whitfield's (1951) that negative and positive information which is formally of the same value is not always recognized as equally useful even by subjects of superior intelligence—this, too, in a simple binary situation. It therefore appears that an explanation other than that of directly given perceptual association must be looked for.

Bruner *et al.*, in their discussion of the use of negative information in concept attainment, make no mention of perceptual links but emphasize, rather, that negative information must usually be transformed into positive information. They speak of the 'inability or unwillingness' of their subjects to use negative information and suggest that information resulting from 'in-the-head' transformation may be distrusted, perhaps because of awareness of the risk of error.

It appeared that some further evidence bearing on these issues might be obtained by a study of matching rather than of concept attainment and, specifically, by having the subjects construct, rather than solve, a matching problem. The construction of a problem calls for an explicit consideration of structural characteristics to a degree not generally necessary in the solution of the problem. It thus provides admirable opportunity for subjects to display their awareness of formal equivalence. (It may also provide them with an opportunity to develop their awareness by obliging them to render explicit that knowledge of structure which often appears to remain implicit in ordinary problem solving. Bartlett and Piaget have both, in recent years, emphasized the importance of awareness of structural characteristics, but the word 'awareness' is ambiguous and Piaget stresses the fact that the subject, in one sense, generally does not know what he is doing. The construction of a problem as opposed to the solving of it may prove to be an excellent way of helping him discover what he is doing and may thus have educational significance, as some teachers realize. It may also be a valuable means of tackling what Bruner *et al.* (1956) describe as 'the great technical problem' in the study of thinking: namely, the externalizing of linked sequences of behaviour for purposes of observing them.) Further, it is possible to impose conditions which make certain ways of proceeding preferable to others, within the context of the task. In the present instance, the advantage of this is that the task can be set in such a way as to place high value on the use of negative information and so give the subject good cause to prefer it to its positive counterpart if he understands the formal equivalence of the two. Suppose n objects $(a, b, \ldots, n)$ to be matched one to one with n other objects $(a', b', \ldots, n')$. Then information necessary for this matching may be given in one of the following ways. (The possibility of using disjunctive information is omitted from consideration.) (i) $n-1$ positive statements, associating a with a', b with b' and so on (information concerning the final association, n to n', being, of course, redundant, since n and n' will be the only objects remaining). (ii) $n-2$ positive statements, associating a with a', etc., and 1 negative statement, dissociating $n-1$ from n' or $(n-1)'$ from n. (iii) $n-3$ positive statements and 3 $[2+1]$ negative statements, which latter may be given in a variety of ways so long as two of them combine to rule out one pairing and the third rules out another. Thus the three negatives dissociating $n-1$ from

n', $n-2$ from n' and $n-2$ from $(n-1)'$ would be adequate for this purpose; but the three negatives dissociating $n-1$ from n', $n-2$ from $(n-1)'$ and n from $(n-2)'$ would not. (The three negatives which are adequate will be described as being in strong combination; the three which are not will be said to be in weak combination.) (iv) $n-4$ positive statements and 6 $[3+2+1]$ negative statements in strong combination. And so on, until the following extreme is reached. (v) $n-n$ positive statements (i.e., none) and

$$[(n-1)+(n-2)+\ldots 1]$$

negative statements in strong combination.

This being so, if each positive piece of information is valued at n points and each negative piece at 1 point, then the total 'score' for an all-negative construction (see (v) above) will be lower than that for any construction containing positives, and subjects can be told of this scoring system and instructed to achieve the lowest possible score.

It can be demonstrated that, with negatives valued at 1 point, n points is the least whole number weighting for positives which makes the all-negative construction the most 'economical'. For, as we have seen, the all-negative model contains

$$[(n-1)+(n-2)+(n-3)+\ldots 1]$$

negatives. Now for each positive that is introduced, the first term of the negatives is omitted: that is, the introduction of 1 positive makes it possible to dispense with $n-1$ negatives; the introduction of a second positive makes it possible to dispense with a further $n-2$ negatives; and so on. If, then, an all-negative construction is to be the most economical, the weighting of the positives must be greater than the greatest number of negatives which a positive can ever replace, namely $n-1$. That is to say, the weighting of the positives must be n.

In the problem used in this inquiry, $n=5$. Each positive was therefore valued at 5 points and the possible ways of providing information were 5 in number, containing 4, 3, 2, 1 and 0 positives respectively. These different possible constructions will be referred to as the $(4, \overline{0})$, $(3, \overline{1})$, $(2, \overline{3})$, $(1, \overline{6})$ and $(0, \overline{10})$ models.

II. PROCEDURE

In the first inquiry, nineteen subjects, aged 14 years to 14 years 6 months, took part. All of them had I.Q.'s within the range 112–118 on each of two Moray House group intelligence tests which had been administered two and three years previously. They had all been taking part in a more extensive investigation which had involved the solution of a matching problem similar to that which they were now to construct. In this earlier problem, information was provided on the $(1, \overline{6})$ model. It must, therefore, be recognized that a set towards this model may have resulted and no reliance can be placed on the evidence of the frequency of its occurrence in this inquiry. However, the subjects' previous work with a matching problem was considered advantageous in so far as it reduced the likelihood of confusions arising merely from a failure to understand what was required. The job of explaining to the subjects what they were to attempt in the problem-construction task was certainly made much easier by this circumstance. The actual wording of the instructions was as follows. (They were read aloud to the child and then the typed sheet was given him to consult.)

Problem A

Do you remember the question about the five boys and the five schools? This time I want you to make up a problem like that instead of solving it. I'll tell you what the answer is to be and you are to tell me what information would have to be given to the children who are to solve the problem. We'll call the boys by their initials. Let's say their names are Arthur, Bob, Charles, Dick and Edward. And we'll just call the schools 1, 2, 3, 4 and 5.

Now here is the answer:

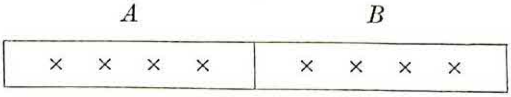
A	*B*	*C*	*D*	*E*
4	2	1	3	5

That is, Arthur goes to school number 4, and so on.

Now you can give information in two ways. Either you can say something like 'Arthur goes to school 4' (which we can write as $A = 4$, for short), or you can say something like 'Arthur does not go to school 2' which we can write as $A \neq 2$, for short). Pieces of information like $A = 4$ we'll call *positive*. Pieces of information like $A \neq 2$ we'll call *negative*. Every positive piece of information that you give will cost 5 points; every negative piece will cost 1 point. You have to try to give enough information for some one who does not know the answer to be able to find it out, and you have to try to do it in such a way as to make the lowest possible score.

When the subjects had made their attempts at construction, they were given a second task which was as follows.

Problem B



Here are two boxes labelled *A* and *B*. The crosses in the boxes stand for the numbers 1–8. You are to find out which numbers are in *A* and which are in *B*, but you do not have to worry about the way in which the numbers are arranged inside the boxes—that is, if you discover that 3 is in *A* you do not have to go on to ask whether it is the second cross or the third one or so on.

You will ask me for the pieces of information you think you will need. You may ask me either for a piece of *positive* information or for a piece of *negative* information; and you may tell me whether you want it to be about box *A* or about box *B*. If you ask me for a piece of positive information about box *A* I shall give it to you in the form:

'One of the ×'s in *A* is 5'—or whatever it may be.

If you ask for negative information about *A* I shall tell you:

'None of the ×'s in *A* is 7'—or whatever it may be.

Now each positive piece of information will cost you 2 points, each negative piece 1 point. You are to try to find the answer making the lowest possible score.

This second problem was suggested by Whitfield's paper (1951) and was introduced as a further check on awareness of the equivalence of positives and negatives. The order of presentation of problems A and B was not varied because it was considered that the number of subjects was too small for any valuable estimate of the effects of differing order to be obtained and that with a group of this size there was more to gain by keeping the conditions of administration constant for all subjects.

Subjects were tested individually and a record was made of all their comments and explanations. They were encouraged to make full use of pencil and paper in working out and checking their constructions, and they were allowed to continue with each problem until they declared themselves satisfied that they had made the lowest possible score. They spent, on the average, 32 minutes on the first problem and 8 minutes on the second one.

In the second inquiry the subjects were ninety women students in their third year at a Scottish teachers' training college. They were tested as a group with problem A, worded in more general terms and without reference to earlier experience of matching problems, which these subjects had not been given. The time allowed was 25 minutes. Problem B was not administered. It could not be given as a group test without radical alteration.

III. RESULTS

(a) *First inquiry—Problem A*

Fourteen subjects were able finally to provide information on the $(0, \overline{10})$ model. No one, however, achieved this at the first attempt. The average number of attempts for all the subjects was 3·4 with a range of from 2 to 8.

Table 1. *First construction attempts*

Col. 1		Col. 2		Col. 3		Col. 4		Col. 5		Col. 6 Total
Model	*f*	Model	*f*	Model	*f*	Model	*f*	Model	*f*	*f*
$4, \bar{0}$	—	—	—	—	—	—	—	—	—	—
$3, \bar{1}$	1	$3, \bar{2}$	1	—	—	—	—	—	—	2
$2, \bar{3}$	2	$2, \bar{4}$	1	$2, \bar{5}$	2	—	—	—	—	5
$1, \bar{6}$	4	$1, \bar{7}$	1	$1, \bar{8}$	1	$1, \bar{9}$	2	$1, \overline{10}$	1	9
$0, \overline{10}$	—	$0, \overline{11}$	1	$0, \overline{16}$	1	$0, \overline{18}$	1	—	—	3

Table 1 shows the construction made by each subject on his first attempt. Models in column 1 are the formally correct variants described in the introduction, though, of course, only $(0, \overline{10})$ is 'correct' in terms of the special conditions imposed. In all the other columns, the models contain redundant information. The models are arranged by rows according to the number of positives in decreasing order from top to bottom of the table, and by columns according to the number of negatives in increasing order from left to right of the table. Thus, the top left corner contains models with most positives and the bottom right corner contains models with most negatives. The cross-totals of frequencies (col. 6) indicate how often 4, 3, 2, 1 and 0 positives respectively were used. Three subjects (models $(1, \bar{7})$, $(2, \bar{4})$ and $(2, \bar{5})$ in columns 2 and 3 of Table 1) began by trying to use only negatives but could not keep this up and introduced positives in the course of their first attempt. No one tried to start by giving four pieces of positive information. All must, therefore, presumably have appreciated in some manner that the use of negatives was advantageous.

The four who began with model $(1, \bar{6})$ proved to be the best subjects. All proceeded to model $(0, \overline{10})$ at the second attempt and made no errors at all, unless their initial inclusion of one positive be termed 'error'. It has already been pointed out that, since the subjects had all had practice in solving a problem constructed on model $(1, \bar{6})$, a set towards it might be anticipated. This would not, however, seem to explain why the four who began in this way all moved very rapidly on to successful completion of the task. What is probable is that some at least of the nine subjects who began by using one positive (see Table 1, fourth row) were influenced by the problem they had solved previously but that only these four were capable of avoiding redundancy from the start. Perfect recollection of the $(1, \bar{6})$ model in the absence of good understanding would seem highly improbable, especially as the names used were changed.

Analysis of the construction attempts showed that there were three main errors.

(i) *Positive/negative redundancy.* This amounts to failure to realize that a positive statement linking two characters makes any negative statement involving either character superfluous. Thus, if $C = 1$, it follows without further statement that $C \neq 2$, that $D \neq 1$, etc. Six children made the error of providing such superfluous negatives.

(ii) *Negative/negative redundancy.* Logically this is barely distinguishable from positive/negative redundancy, since it is only when negatives combine to equal a positive statement that further negatives can be called redundant. Thus, logically, if positive/negative redundancy is understood and avoided one would expect the same to be true of negative/negative redundancy. But this was not found to be the case. Thirteen subjects made errors of negative/negative redundancy, as compared with six in the positive/negative category (four of these being however in both categories), and there was evidence that psychologically the two were not equivalent. Here is an example of negative/negative redundancy associated with explicit awareness of the need to avoid positive/negative redundancy. Words spoken by the experimenter are in brackets.

S. 7. *A* doesn't go to 2.... *A* doesn't go to 3.... *B*.... I am trying to work out whether it would be best to say that *B* goes to 2 or that he doesn't go to any one. (How could you do that?)—Pause—Could I score out '*A* doesn't go to 2' and put '*B* goes to 2'? (Yes. Why did you score that out?) Well, as it is only one boy that goes to a school and you're saying that *B* goes to 2, to say that *A* doesn't go to 2 would be a waste of a point.

Thus this subject shows clearly that he understands positive/negative redundancy. But a little later he goes on:

Well, I see one way to do it but I don't know that it's the lowest of them. You could say *A* doesn't go to 2, 1, 3 or 5, but you would have twenty points. (Why?) Well, as there is one point for every negative you do you would have four points for every letter and there are five letters.

The subject gives the twenty negatives accordingly, but then says he knows this is not the lowest possible score and proposes to reduce it as follows:

Well, you could try saying *A* goes to 4 and you could take away the 4 from *B*, *C*, *D* and *E*.

That is, he proposes to reintroduce a positive so as to be able to reduce the negatives. He shows no awareness that his first four negatives are exactly equivalent to a positive in respect of making such a reduction possible. It is true that, at his last attempt (he needs six in all), he says he 'has a theory that brings it down to 10' and is then able to explain that 'if you have four negatives for the first one you only need three and then two and then one'.

There was no example of awareness of the risk of negative/negative redundancy associated with lack of awareness of the positive/negative variety. The two cases of positive/negative redundancy in the protocols of children who did not make negative/negative redundancy errors arose when difficulty in bringing negatives into strong combination (see (iii) below) had led subjects to the introduction of a positive towards the end of a construction attempt. This positive then rendered earlier negatives redundant and the fact was not noticed by the subjects.

(iii) *Use of negatives in weak combination.* Some of the children showed themselves unable to decide effectively which negative statements would be of most value. For instance:

S. 9. *A* does not go to 2, 1 or 3 and that leaves 4 or 5. And *C* does not go to 4. And *B* does not go to 5. That has told all the ones they don't go to....

Or again:

S. 5. Bob does not go to 1 or 3. Arthur does not go to 2 or 5. Charles does not go.... Bob must go to 2, 4 or 5. Arthur must go to 1, 3 or 4. Charles not to 4 or 5. It's not really telling you anything! So Arthur not to 5 and Charles not to 5. Dick does not go to...1, 4 or 2. So Dick to either 5 or 3. And Bob...I'm getting myself tied up now. (Are you trying to do it all with negatives?) Yes.

This subject is finally able to produce a solution with one positive and ten negatives, that is, with four superfluous negatives.

Errors of this kind were made by seven children.

Six children made a direct statement at one point or another to the effect that, in their opinion, at least one piece of positive information would have to be given; and there seemed to be a very general expectation that 'to do it all with negatives' would be costly because so many would be required.

Here is an example.

S. 11. I think you would have to say one of the schools one of the boys goes to...because if you gave it in negatives I think it would add up to more than five points. (What makes you think that?) Well, you have to give at least four negatives to make one positive, and if you have the positive I think it makes.... If you are giving a reason and you have a positive you don't always have to say the three letters because the three letters or the four letters do not go.

This is certainly not very coherent but the 'at least' is revealing: it should, of course, be 'at most'. It looks, again, as if the subject realizes that a positive can make certain other information superfluous ('you don't always have to say the three letters') but does not recognize that a combination of negatives can do so equally effectively. She seems to be trying to explain that the value of one positive in terms of the knowledge it gives is greater than the value of four negatives. Thus the four negatives 'make' the positive in one sense but do not fully equal it in function.

(b) *First inquiry—Problem B*

Thirteen subjects finally made the lowest possible score by asking for four negative pieces of information relating to the same box. Among the six who failed to do so were four of the five subjects who failed with problem A.

The number of attempts necessary was much less than for problem A. Ten subjects took only one attempt, though among these were four who failed. (It will be recalled that the experiment stopped when the subject said he was sure he had the lowest possible score.) Eight subjects took two attempts and one took three attempts.

In each attempt, it was possible for the subject either to ask for information about one box only or about both boxes; and to ask for negative information or for a mixture of positive and negative information. (Theoretically he could of course also have asked only for positive information but this did not occur.) From Table 2, which shows the numbers in these categories on the first attempt, it will be seen that fourteen of the nineteen subjects have been able to recognize at once, in this situation, the possibility of dispensing with positive information. Only six, however, began by confining their attention to one

Table 2

	One box	Two boxes
Negative only	6	8
Positive and negative	0	5

box alone. The most common tendency was to ask for information about *A* and about *B* in fairly strict alternation.

A number of children tried to look for a principle of distribution: for instance, that all the odd numbers were in *A* and all the even ones in *B*. When this became evident, the subject was at once told that there was no rule to be discovered.

(c) *Second inquiry—problem A*

Of the ninety subjects who took part in the second inquiry: (i) fifty-six provided information adequate for solution and free of redundancy, and of these forty-nine did so using the $(0, \overline{10})$ model, while four used the $(1, \overline{6})$, two the $(2, \overline{3})$ and one the $(3, \overline{1})$ model; (ii) twenty gave more information than necessary; (iii) seven gave less than necessary but considered they had given enough; and (iv) seven papers were unfinished.

Among the twenty-seven finished papers where information was either redundant or inadequate—(ii) and (iii) above—there were twenty-three instances of negative/negative redundancy, three of positive/negative redundancy and one of negatives in weak combination. (Notice that the fact of giving less information than is necessary for solution in no way rules out the possibility of giving redundant information: in what information *is* provided there may be duplication.)

Scrutiny of the fifty-six papers in (i) above showed that, in many cases, earlier attempts had been written down. Some of these had obviously been abandoned mid-way and sometimes two or three attempts had been scored out in such a way that it was hard to tell where one ended and the other began. Thus, no exact count of errors in these earlier attempts could be made; but it seemed that, on the whole, the proportion of positive/negative to negative/negative redundancy obtained by analysis of the twenty-seven papers in groups (ii) and (iii) above was maintained.

IV. Discussion and conclusions

The finding that logically equivalent positive and negative information is not always psychologically equivalent is confirmed in this inquiry. It is confirmed with particular force by virtue of the fact that the problem situations were not neutral with respect to positives and negatives. Rather the conditions of the tasks were such as to give the negatives privileged status and encourage the subjects to work with them. Evidence of 'inability or unwillingness' to do so is, then, in these circumstances, all the more strong.

There is some unambiguous evidence of inability, as opposed to unwillingness. This is to be found in the fact that three subjects began by trying to produce an all-negative construction for problem A, found themselves in difficulties and introduced positives to simplify their task. It is certainly true that the use of negatives calls for an appreciation of the structure of a matching problem in a way that the use of positives does not. Lack of this appreciation shows itself in the failure to bring negatives successfully into strong combination and, when this happens, resort to positives is only to be expected. But this sort of difficulty with the structure of the problem cannot account for all the evidence of lack of equivalence between positive and negative information. In particular, it cannot explain the evidence that positive/negative and negative/negative redundancy are psychologically distinct—evidence obtained both from the individual studies (see S. 7 and S. 11) and from the group inquiry which showed that, with older subjects, one of these errors had largely disappeared while the other continued to be extremely prevalent and, indeed, to be the only common source of trouble remaining, since the difficulty of understanding how to bring negatives into strong combination had also diminished to vanishing point.

How, then, are we to interpret the fact that it is possible to understand quite well the way in which negatives can 'add up' to 'equal' a positive, yet not see that to provide negative information which duplicates this derived positive is to introduce redundancy as surely as if one provides negative information which duplicates a directly stated positive?

The suggestion made by Bruner *et al.* (1956) that the need for 'in-the-head' transformation explains the peculiar difficulty of negatives will not serve for this purpose. In the matching problem construction task, repeated transformation in both directions was called for. What we have to explain is the fact that one particular sequence of transformations—from negative to positive and back again—is especially difficult, a 'block' of some kind occurring commonly after the first transformation has been made. Nothing of the kind was ever observed to occur when the opposite sequence—from positive to negative and back to positive—was in question. No subject showed any sign of regarding negative information that had been derived from positive information as in any way different in status or usefulness from directly given negative information. The possibilities for transforming positives into negatives were not always fully exploited—witness the occurrence of positive/negative redundancy—but once this transformation had been carried out there was no suggestion that the derived negatives could not be further utilized.

There appear to be two possible explanations of failure to use the derived positive. One is that negative information is in some way distrusted. However, the distrust would not, in this case, seem to stem from awareness of the risk of error in the course of transformation, as Bruner *et al.* suggested it might do, but rather from a less rational feeling that negative information is not such good currency as positive information and that complicated and far-reaching transactions cannot safely be made with it. There is, indeed, occasional evidence in the protocols that negative information is distrusted even when no transformation is involved. For instance, S. 2 comments '*E* doesn't go to 4 and *A* does, so *E* can't be 4'. Here a directly given negative is reinforced by a positive before it is accepted as quite conclusive.

Another possibility, however, which must not be overlooked in considering the limited transformability of negatives is that this particular manifestation of lack of equivalence is to be explained by reference to the attitude of the subjects to positive information. It may be due, in part, to a feeling of finality attaching to positive statements and deriving from the knowledge that positive matchings are the ultimate end of the endeavour. What may be lacking is the ability to see them as ends in themselves and yet at the same time as means to further ends within the framework of the problem.

The results of problem B are, on the whole, easier to interpret than those of problem A. As Table 2 shows, most of the subjects were able to recognize at once the formal equivalence of positive and negative information in this simpler situation. But it should, of course, be noted that they had all just previously tried to solve problem A. That even after this experience five subjects still included positives—and one of them went so far as to state explicitly that at least one positive was absolutely necessary for the solution of B—is quite impressive evidence of unwillingness to rely on negatives alone. In this case, it does seem appropriate to talk of unwillingness rather than inability, because no subject in problem B relied entirely on positives and, in this problem, to understand how one piece of negative information can be used to assign a number positively to the other box might be said to be evidence of 'ability' to assign all numbers in this way. The insistence on retaining one positive appears, then, as a security measure rather than a failure of understanding.

The greatest source of difficulty in problem B lay in understanding the importance of asking for information about one box only. This, of course, can be regarded as yet another instance of difficulty with negatives—or at least with indirect information—since it depends on failure to realize that to know the complete contents of one box is to know where the remaining numbers are not located, which, in this binary situation, is to know where they are located—that is, to know the complete contents of both boxes.

It should be noticed that the tendency for subjects to ask about boxes *A* and *B* in systematic rotation was less likely to lead to rapid success than a more haphazard procedure whereby there was a chance that all spaces in one or other box would be allotted before seven pieces of information had been obtained. When this happened, the subject generally realized that he had all the information he needed, and this realization, coming when more than one of the eight spaces remained unfilled, seemed to facilitate the understanding that only four pieces of information were necessary.

The tendency to look for a rule or principle was probably an obstacle in the way of this crucial understanding, for the subject who had a hypothesis presumably based his decisions about which box to choose on irrelevant considerations relating to the testing of the hypothesis. 'Looking for a rule' in this situation perhaps exemplifies a general tendency observed by Bruner *et al.*: the tendency for subjects to feel that they must, on pain of failure, try to anticipate trends, rather than wait for what the data reveal.

I am grateful to the Edinburgh schools concerned and to Moray House College of Education for the facilities which they provided.

References

Bruner, J. S., Goodnow, J. J. & Austin, G. A. (1956). *A Study of Thinking*. New York: Wiley.

Hovland, C. I. & Weiss, W. (1953). Transmission of information concerning concepts through positive and negative instances. *J. Exp. Psychol.* **45**, 175–82.

Smoke, K. L. (1933). Negative instances in concept learning. *J. Exp. Psychol.* **16**, 583–8.

Whitfield, J. W. (1951). An experiment in problem solving. *Quart. J. Exp. Psychol.* **3**, 184–97.

(*Manuscript received* 2 *May* 1958)

AN INTERPRETATION OF THE 'DISPLACEMENT' PHENOMENON*

By DALBIR BINDRA

McGill University, Montreal, Canada

The observations that animals obstructed in the execution of a particular ongoing or customary activity (e.g. attacking, copulating) tend either to direct the same activity toward another object or to engage in a completely different activity (e.g. doodling, grooming, preening) have traditionally been described as 'displacement phenomena'. Both psychologists and ethologists currently interpret these phenomena in terms of some kind of *displacement mechanism*, which is assumed to displace the 'energy' or 'drive' from one reaction system to another. This type of interpretation is vague and *ad hoc*, and at best provides only a redundant description of the observed phenomena. The alternative interpretation proposed here looks upon the problem of the occurrence of the so-called displacement activities as a special case of the general question of the factors determining the occurrence of *any* activity that exists in an animal's repertoire. In particular, it is suggested that all instances of displacement phenomena can be adequately accounted for in terms of the operation of three factors: (*a*) an increase in the level of arousal of the animal brought about by the obstructing event, (*b*) the relative habit strengths of the various activities in the repertoire of the animal, and (*c*) the nature of the sensory cues provided by the altered stimulus situation. This formulation incorporates the important ethological and psychological findings, and provides an interpretation of displacement phenomena which is not *ad hoc* and which suggests new lines of experimental work.

I. THE PHENOMENON AND ITS CURRENT INTERPRETATION

When an animal is prevented from engaging in a particular activity, either by an environmental obstacle or by interference from some other ('conflicting') responses, that activity may be said to be obstructed. Such obstruction ('frustration') of an on-going, customary, or in some sense 'expected' activity affects the course of the animal's subsequent behaviour. A sequence of events of this kind, response A–obstruction–response B, has been widely described as *displacement phenomenon*, and the subsequent activity (response B) is labelled *displacement activity*.

Generally speaking, the term 'displacement' is used to describe two different types of consequences of the obstruction of an activity (or 'drive'). The first type includes cases in which, when an opportunity arises, the animal engages in the same general class of activity as the one that was obstructed, but the activity is directed towards a different object. Thus, a man who, after some provocation, is prevented from hitting his boss because of some interfering 'respect responses', later, may (at least according to folklore) act aggressively toward his wife or children. Secondly, the term 'displacement', is employed to describe phenomena in which, following the obstruction, the animal shows some other, 'irrelevant', but fairly specific, activity. Thus, a child prevented from eating candy may start to suck his thumb, or a sheep prevented from escaping a noxious stimulation may begin to graze. In cases of this type, the animal does not engage in the activity that was originally obstructed, but in some completely different activity. Ethologists (e.g. Lorenz, 1935, 1941; Tinbergen, 1951, 1952) employ the term 'displacement' in this sense when, for example, they state that, if conflicting responses obstruct fighting in a skylark, or copulation in a duck, the animal begins to preen or peck intermittently. The various behavioural phenomena described by such labels as 'substitution', 'sublimation', and 'displaced aggression' are all special cases of the two types of consequence of obstruction described above.

---

* Based on a talk given in a symposium on motivation held at the annual meeting of the Canadian Psychological Association, Edmonton, June 1958.

The currently popular interpretation of these phenomena, in both psychological and ethological writings, is in terms of a *displacement mechanism* (Freud, 1913). Thus, it is thought that, when an animal is prevented from engaging in an activity (or 'expressing a drive'), the particular response tendency or its accompanying 'energy' remains active until it can be dissipated either by 'displacing' it on to an irrelevant object or into an irrelevant activity. As Tinbergen has put it: 'Our hypothesis therefore is that the displacement activities are outlets through which the thwarted drives can express themselves in motion' (1952, p. 12). Such a notion of a displacement mechanism is quite inadequate to explain displacement phenomena. It is vague and *ad hoc*. When an animal is prevented from engaging in an activity for which it is ready, it obviously must show some activity—if it is still alive. Only two possibilities exist. Either its new activity will resemble, to a lesser or greater extent, the prevented activity, or the new activity will be completely different. In neither case is anything gained by stating, *post hoc*, that it represents a 'displacement' of the original response tendency, energy, or drive. Invoking this hypothetical mechanism so arbitrarily provides only redundant descriptions and is unproductive of experimental analysis. In order to formulate a meaningful and experimentally fruitful hypothesis, the exact empirical variables that control the occurrence of displacement activities must be stated explicitly. (Actually, to describe the phenomena under discussion as *displacement* phenomena is to beg the question; however, in order to avoid coining a new term, I shall continue employing the term 'displacement phenomena' as a descriptive label, without accepting the 'displacement mechanism' as an adequate explanation.)

A number of writers have already confronted the difficult task of specifying the empirical variables that determine the occurrence of displacement activities. For example, Tinbergen (1952) has discussed the roles of factors such as initial posture and external stimulation, and Andrew (1956*a*, *b*) of stimuli originating from the body surface. Miller (1948) has experimentally analysed some of the factors operating in 'displaced aggression'. In the following pages an attempt is made to incorporate the findings of ethological and psychological workers into a general framework which would present an advance over the displacement notion. The proposed framework appears to have two advantages. First, it suggests new lines of experimental work. Secondly, it is not *ad hoc*; rather, it is a special application of the same general basic concepts as are employed in accounting for other types of behaviour.

II. An alternative interpretation

The general point of view adopted here is the one I have more fully developed elsewhere (Bindra, 1959). Basically, it assumes that the problem of the analysis of any particular type of behaviour can be reduced to two (empirical) questions. First, how did the particular activity develop in the organism's repertoire? Secondly, what factors determine that this activity, rather than any other activity of which the organism is capable, will occur at a given time and place, and the details (e.g. latency, duration, errors) of the way in which the activity occurs? In the present analysis of displacement activity, the first question can be ignored; we can assume that the activities usually described as displacement activities (preening in birds, grazing in sheep, grooming in chimpanzees, thumb-sucking in the human infant, and so on) have already developed in the organism's repertoire. (The exact roles of innate and experiential factors in the development of these activities still remain to be determined: see Bindra, 1959, chapters 3 and 4.) Thus, the question under discussion is one of the factors that determine whether or to what extent will any given activity occur at a particular time and place. In general, as I have argued in the reference cited, four main sets of factors completely determine the occurrence of *any* activity: habit strength of the activity, sensory cues provided by the stimulus situation, level of arousal of the organism, and state of blood chemistry of the organism. In this discussion an attempt is made to account for the occurrence of the so-called displacement activities in terms of the operation of the first three of these factors. The discussion involves a number of general propositions and a consideration of their application to displacement phenomena.

(1) *The level of arousal of an organism is raised whenever it is exposed to an environmental change or novel sensory stimulation.* (The term 'level of arousal' is employed here to denote roughly the degree to which an organism is excited rather than calm. The low extreme end of this dimension is represented by the state of deep sleep or general anaesthesia, the high extreme end by the state of panic or epileptic seizure. The commonly employed indices of arousal include such measures of physiological function as electrical resistance of skin, muscle action potentials, and the pattern of firing of nerve cells within the brain. For a fuller discussion of this concept, see Duffy (1951, 1957), Lindsley (1951), Schlosberg (1954), Hebb (1955), Malmo (1957) and Bindra (1959).)

The extent to which arousal level is raised depends upon the exact nature of the stimulus change; for example, noxious stimulation (such as electric shock or its threat) and socially insulting events tend, in general, to raise arousal level more than does the presentation of lights, noises and simple tasks. Now, the starting-point in the description of the typical displacement phenomenon is some type of obstruction that is described as frustrating, conflict-provoking, anxiety-producing, stressful, or a situation of 'great "tension"' (Tinbergen, 1952). Whatever the exact nature of these situations, they all involve a change in sensory stimulation of the type that substantially raises the level of arousal ('excess of drive'—Tinbergen), though the degree of increase in arousal varies from situation to situation and individual to individual. Thus, the so-called displacement activities typically occur when the animal is in a state of heightened arousal. The heightened arousal level typically lasts for some time after the cessation of the event (i.e. change in sensory stimulation) that initially produced it. According to the present view, it is the continued presence of the heightened arousal level that is partly responsible for what the animal does subsequently. In the present context, this means that the particular (displacement) activities that occur in a given situation depend upon the degree and duration of the heightened arousal level.

(2) *A shift in the level of arousal of an animal alters the relative probabilities of occurrence of the various activities that exist in its repertoire.* At very high (and very low) levels of arousal there is an increase in the relative probabilities of occurrence of two types of activities: (*a*) activities that were initially acquired or recently practiced at a high arousal

level, and (*b*) activities that have a high habit strength, that is, those that have been frequently practised. It is the high habit-strength activities that are particularly relevant here. According to the present view, the activities that are usually referred to as displacement activities are those that are highly prepotent in the repertoire of organisms. As Tinbergen has stated, 'an animal is apt to use as outlets those patterns which for some reason offer least resistance' (1952, p. 18). An enumeration of the various displacement activities indicates that they are typically high habit-strength activities. For example, thumb-sucking in children, grazing in sheep, preening in birds are activities that are likely to be over-practised in the repertoire of the members of the particular species.

The above proposition states that activities not prepotent in the repertoire of an animal are not likely to occur when it is frustrated, made anxious, put in a conflict situation, or aroused in any other way. Thus, a man in whom aggressive acts are not prepotent is less likely to act aggressively in *any* situation than one in whom aggressive acts have reached a high level of habit strength. This statement implies that the man who, on being reprimanded and, thus, provoked by his boss, does not act aggressively in the office situation is not likely to act aggressively even when he sees his wife or children. But the man who does act aggressively in the office situation is also likely to act aggressively in the home situation, provided his arousal level remains high. These predictions are directly opposed to those that might be made on the basis of the traditional interpretation in terms of a displacement mechanism. That would predict that an individual who is provoked but does not act aggressively ('suppresses his anger') in the office situation is more likely to act aggressively ('let go') in the absence of his boss. These opposing predictions are subject to an empirical test. In this connexion it is interesting to note that, in Miller's (1948) experiment, rats acted aggressively toward a doll after they had obtained considerable practice in attacking each other in the experimental situation. The present interpretation would predict that aggression toward the doll would occur only in the case of those rats in whom aggressive acts had acquired a certain (undertermined) degree of prepotence. Similarly, according to our formulation, birds that normally preen more than others will also be more likely to show preening when they are aroused than birds in whom preening has not reached the same level of habit strength. These predictions can be tested experimentally. The present suggestion of interaction between the factors of habit strength and arousal makes it unnecessary to postulate different (undetermined) 'thresholds' (Tinbergen, 1952) for the occurrence of various (displacement) activities existing in the animal's repertoire.

(3) *Any activity in the repertoire of animals can occur only within a certain range of variation of the sensory cues with which that activity is associated; when such sensory cues are altered beyond a certain degree that activity cannot occur at all.* It is well known that if the situation in which an animal normally eats or copulates is altered beyond a certain degree, the animal will cease, at least for some time, to engage in the customary activity; similarly, an animal which acts aggressively toward a weaker animal may act submissively toward a dominant one. In the present context, this proposition means that, when an animal has been aroused in some way, its subsequent behaviour will depend upon the nature of the sensory cues generated by the situation in which it finds itself. If, for example, after some frustration, a rat is placed in the situation in which it normally copulates, then the rat is likely to copulate; if it is placed in an 'attack situation', it will tend to show aggressive behaviour; and if it is placed in an 'eating situation', it will probably eat. In short, the

sensory cues associated with a particular activity will evoke that activity, except in so far as the customary activity may be affected by the raised level of arousal. Thus, according to the present view, the behaviour following any obstruction is determined by the sensory cues *currently* active in the animal, and not by any residual 'energy' made superfluous by the obstruction. The influence of initial posture and environmental stimuli discussed by Tinbergen (1952) is considered here as a special case of the control on behaviour exerted by the current (proprioceptive and exteroceptive) sensory cues. Where only some of the relevant sensory cues are present, an activity may occur only incompletely; thus, while normally a bird may peck and pick up food, if no food is in sight it may peck and pick up nothing or may interrupt a peck in progress.

The above proposition implies that activities associated with more frequently occurring sensory cues will be more likely to occur following obstruction than activities associated with less frequently occurring cues. Whereas most activities of animals are dependent upon appropriate environmental cues, there are some that seem to be connected primarily to cues arising from the animal's own body (Andrew, 1956*b*). Activities of this latter type include grooming in rats, preening in birds, thumb-sucking in infants, and tics and stereotyped postural movements in most animals. If the present proposition is accepted as true, it is not surprising to note that it is activities of this type that are most frequently described as displacement activities. The proposition also suggests a way to account for the observation that, for example, a man provoked by his boss may not attack his boss but may, a minute later, bang his own desk or reprimand his own secretary. In such a case it is likely (and this can be determined independently of any particular incident) that the boss-in-his-office situation does not provide for this man the cues for aggressive acts, but his own secretary and desk do. Similarly, according to the present view, on reaching home the man will act aggressively toward his wife only if that is one of the activities for which the wife is a cue; if his wife is a cue for relaxing, he may 'cry on her shoulder' (and this probably happens more frequently than the current folklore leads one to suspect).

III. CONCLUSION

The interpretation of displacement phenomena proposed here looks upon all such phenomena as resulting from the operation of three general sets of variables: level of arousal, habit strength and sensory cues. This formulation is not considered as, in some sense, a final explanation of displacement phenomena. Rather, it is offered as a line of speculation that points to the type of research that is likely to lead to an adequate explanation. This approach is an advance over the traditional interpretation in two ways. First, it is not *ad hoc*; rather it is a special application of the same basic concepts that are generally employed in accounting for all types of behaviour. Secondly, it appears to be more fruitful in suggesting new lines of experimental work; in particular, the exact effects of the three sets of variables on the occurrence of various types of (displacement) activities emerge as important experimental problems.

REFERENCES

ANDREW, R. J. (1956*a*). Some remarks on behaviour in conflict situations, with special reference to *Emberiza* spp. *Brit. J. Anim. Behav.* **4**, 41–5.

ANDREW, R. J. (1956*b*). Normal and irrelevant toilet behaviour in *Emberiza* spp. *Brit. J. anim. Behav.* **4**, 85–91.

BINDRA, D. (1959). *Motivation: a Systematic Reinterpretation.* New York: Ronald.

DUFFY, E. (1951). The concept of energy mobilization. *Psychol. Rev.* **58**, 30–40.

DUFFY, E. (1957). The psychological significance of the concept of 'arousal' or 'activation'. *Psychol. Rev.* **64**, 265–75.

FREUD. S. (1913). *The Interpretation of Dreams.* London: Allen.

HEBB, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.* **62**, 243–54.

LINDSLEY, D. B. (1951). Emotion. In S. S. Stevens (ed.), *Handbook of Experimental Psychology.* New York: Wiley.

LORENZ, K. (1935). Der Kumpan in der Umwelt des Vogels. *J. Ornithol.* **83**, 137–213, 289–413.

LORENZ, K. (1941). Vergleichende Bewegungsstudien an Anatinen. *J. Ornithol.* **89**, 194–294.

MALMO, R. B. (1957). Anxiety and behavioral arousal. *Psychol. Rev.* **64**, 276–87.

MILLER, N. E. (1948). Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. Abnorm. Soc. Psychol.* **43**, 155–78.

SCHLOSBERG, H. (1954). Three dimensions of emotion. *Psychol. Rev.* **61**, 81–8.

TINBERGEN, N. (1951). *The Study of Instinct.* London: Oxford University Press.

TINBERGEN, N. (1952). 'Derived' activities; their causation, biological significance, origin, and emancipation during evolution. *Quart. Rev. Biol.* **27**, 1–32.

(*Manuscript received* 29 *May* 1958)

THE PAIRED USE OF PROJECTIVE STIMULI*

By JOHN LIGGETT

Department of Psychological Medicine, King's College,
University of Durham

One way of avoiding some of the verbal and other difficulties of projective testing might be to use several stimuli simultaneously. Comparisons of rankings (or other arrangements) would allow inferences to be made about the determinants of judgement. An inquiry is described in which seven vague 'portraits' were sorted (by paired comparison) for a number of different personality qualities. Agreements and disagreements between these sortings were then computed by a brief graphical method (which also yielded a measure of the consistency with which choices had been made). Normal and psychiatric subjects appeared to differ in the importance they attached to certain qualities of personality and in the consistence with which they made their judgements.

I. Introduction

Poorly structured stimulus materials have long attracted the attention of students of personality and motivation. In spite, however, of the apparent sensitivity of these materials and their ability to evoke such wide varieties of response, attempts at rigorous analysis of test products have often been disappointing. This may be due, in part, to the heavy reliance which is placed, in projective testing, on the spontaneous verbal contribution of subjects of very varied linguistic ability. Another difficulty is that the extent of the subject's involvement in the task is difficult to assess and it may not always be fair to assume that he is free from conscious or unconscious restraints on the free expression of private feelings. If these arguments are correct, then much might be gained if means could be found for reducing or abolishing the verbal demands made on the subject and if a method could be found for presenting stimuli in some indirect way so that the subject could feel relatively uncommitted by his responses. It would be helpful, too, if some index of co-operation could be established.

One way of approaching these objectives would be to employ several unstructured stimuli simultaneously and to seek directional or relative judgements between all of them taken together rather than isolated verbal judgements of single stimuli. Comparisons between different arrangements (for example, rankings) of the same stimuli by the same subject might then throw light on the determinants of judgement. For example, the constituents of some general over-all judgement of a number of stimuli might be explored by comparing the ranking 'representing' it with other rankings 'representing' much more specific and closely defined qualities. The relative contribution made by these 'specifics' to the over-all judgement might then be assessed. (An example of this method using vague human portraits will be described later.)

As Kendall (1955) has pointed out, however, circumstances may arise in which the ranking task is an impossible one (notwithstanding the willingness and apparent success of the subject) and it would, in all cases, be necessary to show that choices were not random, i.e. that some consistent principle was, in fact, determining choice of rank position. Such consistency would need to be complete if the rank difference correlation method were to be employed. To achieve any significant analysis of the determinants and consistency of judgement it seems likely that the method of paired comparison must be used and Kendall (1955) proposes two measures based on paired comparison data which are adaptable to our problem. (1) A coefficient of consistence (ζ) which expresses the non-random character of a single set of paired comparison choices, i.e. whether and to what degree a definite principle is governing choice. (2) A coefficient of agreement (u) between two sets of paired comparison data which is free of the difficulty of ρ (viz. that consistence must be complete for correlation by the rank difference method).

---

* Based on papers read at the XVth International Congress of Psychology, Brussels, 1957, and at the Annual Meeting of the British Psychological Society, 1958.

It becomes apparent that the practical usefulness of the procedure proposed will be governed by the extent to which it is possible to reduce the routine labour of administration, and of computation of measures of rank similarity and the precautionary measures of consistency.

In the account which follows a brief graphical method is presented which enables four paired comparison sortings of seven stimuli to be made and all resulting measures of rank similarity and consistency to be computed in a total time (administration and scoring) of a quarter hour. It should be emphasized that these two coefficients are regarded, throughout this inquiry, simply as test scores and not as summary statistics in their own right. The particular inquiry reported here (an attempt to deal with problems raised by McLeod (1947), Allport (1943), Foulds (1948), Hilgard (1949), and Cronbach (1955)) was an attempt to compare attitudes to certain personality qualities ('nervousness' and 'dominance') held by normal and psychiatric subjects—and to assess the tendency of these two groups to attribute such qualities to themselves. The stimulus materials employed were vague, diffused human portraits. At the suggestion of G. W. Allport (personal communication, 1957), a number of precautions were observed in the selection of these portraits, e.g. they did not differ markedly in their intrinsic popularity and none evoked obviously stereotyped responses. The mode of selection and construction of the 'portraits' is described later. The first objective of this inquiry was to have the subject order the portraits into a sequence which represented his conception of what was good and what was bad in human nature. Such an arrangement would, it was hoped, arise from the subject's quite unique perception of the 'qualities' of the people depicted and his private conception of the relative worth of these qualities. The extent to which the perceptions *were* unique would, no doubt, be increased by their unclearness. To extract the information implicit in this ordering it was then proposed to have the portraits ranked again—this time for some much more specific human quality—such as 'dominance'.

A high correlation of the 'best person' order and this 'dominant' order would then suggest a high valuation of dominance. A comparison of 'like self' and 'dominant' orders might indicate the self-view in relation to this particular quality, and a comparison of 'self' and 'best' orders something of the over-all valuation of the self in terms of just those diffusely-felt, unspoken qualities which happened to be held in high regard.

II. METHOD

(*a*) *Materials* consisted of two identical piles of seven diffused distorted portraits, of the same sex as the subject, spiral-bound side by side to facilitate paired comparison (Liggett, 1957*a*). The stimuli were made by optically diffusing and distorting the image during photographic enlargement. The portraits were selected as follows. For each sex, a miscellaneous collection of 40 portraits was shown to 50 normals, neurotics and psychotics for their informal comment. Portraits evoking stereotyped and popular responses were discarded until 20 were left. These were diffused and distorted and shown to 50 normals for their comment. Again, those evoking popular or stereotyped responses were removed until 7 of apparently equal popularity and visibility remained.

(*b*) *Subjects.* One hundred and sixty adult subjects between 20 and 60 years were employed. Eighty had been referred for psychiatric outpatient treatment but were unselected by diagnosis ('Psychiatric group') and 80 were unpaid volunteers (students, cleaners, porters, nurses) who had never sought or been given psychiatric treatment ('Normal group'). There were equal numbers of males and females, of 'intelligent' (at or above average) and 'unintelligent' (below average Raven's Matrices 1938), and of 'old' (41–60) and 'young' (20–40). There were thus 16 subgroups of 10 subjects each.

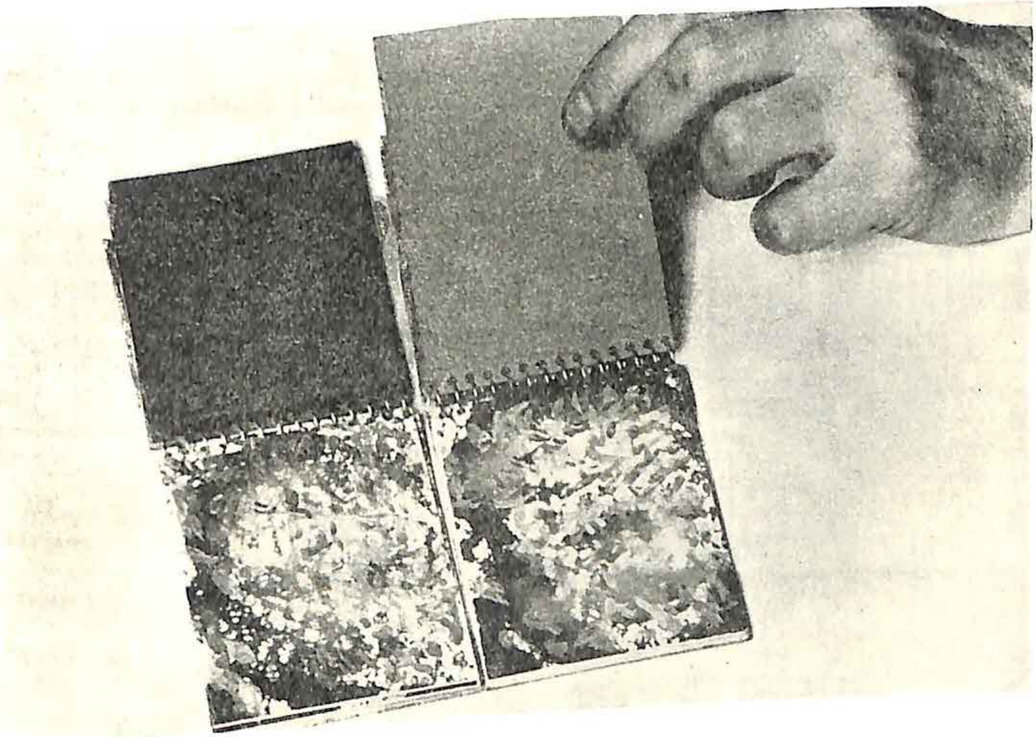


Fig. 1

(*c*) *Administration.* The test booklet (Fig. 1) was placed on the table about 2 ft. from the subject, face 1 on the left and face 2 on the right exposed, and the subject asked 'Which of these two is probably the better person?' Then 1 and 3 were compared and so on until 21 comparisons had been made.

The portraits were then turned back to the beginning, the subject reassured that he was doing well, and then asked, as 1 and 2 were exposed again, 'Which of these two is the more dominant person, I mean who would probably get his own way?' Then, 'Which of these two is probably the more nervous and highly strung?' Finally: 'Which of these two is probably most like yourself *as a person*. It's your first impression that's wanted so don't spend too long.' Verbal comments on the portraits were then invited.

III. SCORING AND RESULTS

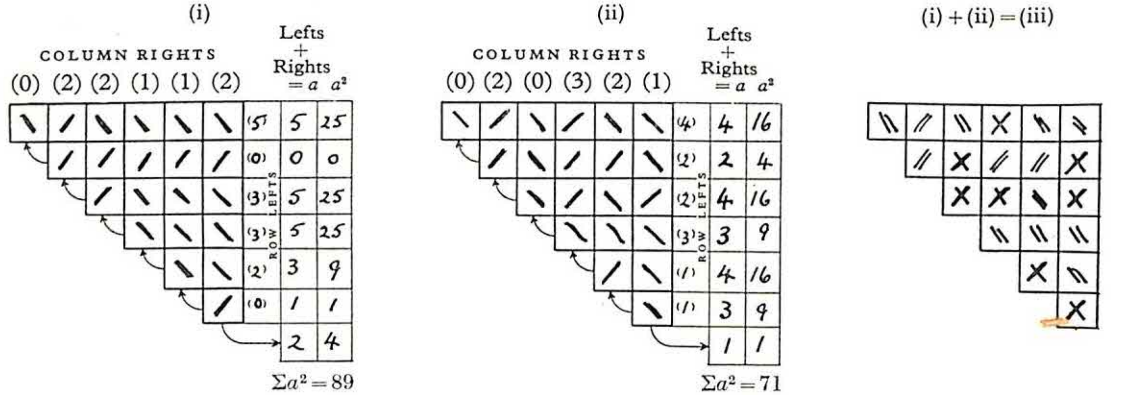
(*a*) *Agreement between sortings.* Choices were recorded by a mark ╲ in one of the 21 cells of a paired comparison framework if the left-hand portrait of a pair was chosen and by a mark ╱ if the right-hand portrait was chosen. The reason for this particular convention was that disagreements between two sets of paired comparison data would later become apparent as X when scoring sheets (which were printed on translucent paper—Liggett, 1957*b*) were superimposed (see Fig. 2). Kendall's Coefficient of Agreement (u) is defined by the expression

$$u = \frac{2\Sigma}{\frac{1}{2}n(n-1)} - 1,$$

where Σ is the sum of the number of cell agreements in a full diagram between two sortings and n the number of objects compared. For the particular circumstances of this inquiry where $n = 7$ and the portraits were compared with each other only once (i.e. 21 comparisons) the coefficient of agreement is reducible to the form

$$u = 1 - \frac{2\Delta}{21},$$

where Δ is the total number of cell *disagreements* in a half diagram. A simple scale may therefore be prepared which converts Δ (the number of cell disagreements) into u (the Coefficient of Agreement).



$$\zeta = 1 - \frac{91 - \Sigma a^2}{28} = 1 - \frac{91 - 89}{28}$$

$= 0{\cdot}93 =$ coeff. of consistence.

$$\zeta = 1 - \frac{91 - \Sigma a^2}{28} = 1 - \frac{91 - 71}{28}$$

$= 0{\cdot}28 =$ coeff. of consistence.

Δ (No. of $\times$'s) $= 8$

$$\therefore u = 1 - \frac{2\Delta}{21} = +0{\cdot}24$$

= coeff. of agreement.

(i) An almost perfectly consistent set of choices (a neurotic's 'Dominance' choices)

(ii) A highly inconsistent set of choices (fictitious example)

(iii) The result of superimposition of translucent sheets (i) and (ii)

Fig. 2. Method for computing 'agreement' and 'consistence'.

An example illustrating the 'X-method' is shown in Fig. 2 in which the right-hand diagram (iii) shows the effect of the superimposition of diagrams (i) and (ii). It will be seen that there are 8 disagreements showing as crosses, that is, $\Delta = 8$ (i.e. $u = +0{\cdot}24$).

Table 1 shows, for each pair of sortings, the number of subjects having 0–5 disagreements, those having 6–11, and those having 12–21 disagreements between two sortings. These particular limits were chosen because they enclosed those cases having a coefficient of agreement (u) between $+1{\cdot}0$ and $+0{\cdot}5$, those between $+0{\cdot}5$ and 0 and those between 0 and $-1{\cdot}0$. Inspection of Table 1 and of the values of χ^2 suggests that the normal subjects tended to attribute dominance more and nervousness less to themselves than did the psychiatric group. Again the normal subjects tended more to 'value' dominance than did the psychiatric group. There was no difference in the 'self/best' comparisons between the two groups and the greater tendency of the psychiatric group to value nervousness was barely significant.

(*b*) *Consistency of sortings.* Kendall's Coefficient of Consistence (ζ) for a single set of paired comparison data is defined by the expression:

$$\zeta = 1 - \frac{24d}{n^3 - n},$$

where n is the (odd) number of objects compared and d the number of 'inconsistent triads' (Kendall, 1955, p. 146). Now

$$d = \tfrac{1}{12}\, n(n-1)\,(2n-1) - \tfrac{1}{2}\Sigma a^2,$$

where a is the row total in a complete paired comparison framework of the type given by Kendall (1955, p. 145), therefore

$$\zeta = 1 - \frac{91 - \Sigma a^2}{28}.$$

Table 1. *Number of disagreements between two sortings*

		Disagreements				
		0–5	6–11	12–21	χ^2	P
Self/best	Normal	62	14	4	4·39	>0.05
	Psychiatric	50	22	8		
Self/dominant	Normal	42	24	14	16·02	<0.001
	Psychiatric	22	22	36		
Self/nervous	Normal	6	11	63	30·50	<0.001
	Psychiatric	26	25	29		
Dominant/best	Normal	37	29	14	10·91	<0.01
	Psychiatric	18	36	26		
Nervousness/best	Normal	7	26	47	6·42	<0.05
	Psychiatric	11	38	31		

The row total 'a' may be computed from a diagram of the type following the 'lefts' and 'rights' convention described earlier by totalling the 'row lefts' in a complete diagram of 42 comparisons or by totalling 'row lefts' plus associated 'column rights' in a half-diagram of 21 comparisons.

Fig. 2 illustrates the method of computing Σa^2 from a half-diagram (21 choices). In diagram (i) a highly consistent set of choices are shown. In diagram (ii), on the other hand, a highly inconsistent set of choices are shown. The formula shows that consistence is a maximum ($\zeta = 1{\cdot}0$) when $\Sigma a^2 = 91$ and a minimum ($\zeta = 0$) when $\Sigma a^2 = 63$. A simple scale may therefore be prepared for converting Σa^2 into ζ the coefficient of consistence.

Inspection of Kendall's Appendix Table 9 (1955) shows that for a 0·95 probability level d must be 4 or fewer.

Now

$$\zeta = 1 - \frac{24d}{n^3 - n} = 1 - \frac{d}{14}.$$

Thus for significance

$$\zeta \text{ must be greater than } 1 - \tfrac{4}{14} = 0{\cdot}714.$$

Values of ζ in excess of 0·714 may therefore be said to show significant consistence. (*N.B.* It was not, in practice, necessary to compute ζ in each case since by substituting for $d = 4$ (and $n = 7$) in the formula

$$d = \tfrac{1}{12} n(n-1)\,(2n-1) - \tfrac{1}{2}\Sigma a^2,$$

Σa^2 assumed the value 83. Cases in which Σa^2 was equal to or greater than 83 could therefore be said to show significant consistence.)

Table 2 shows, for each quality and for each group, the number of subjects achieving significant consistence and the number of subjects not achieving significant consistence. In only one type of sorting (for 'better person') did a difference appear between normal and psychiatric groups.

Table 2. *Consistency of single sortings*

		$\zeta > 0·714$	$\zeta < 0·714$	χ^2	P
Best	Normal	72	8	4·45	$< 0·05$
	Psychiatric	61	19		
Dominant	Normal	60	20	1·65	$> 0·05$
	Psychiatric	52	28		
Nervousness	Normal	56	24	1·744	$> 0·05$
	Psychiatric	47	33		
Self	Normal	71	9	2·25	$> 0·05$
	Psychiatric	63	17		

(*c*) *Retest results.* All of the males were retested for the single quality 'Best' between 2 and 4 months after first testing. The mean value of test-retest disagreements was 2·4 with S.D. 0·41. This corresponds to a mean coefficient of agreement between test and retest of approximately +0·7.

(*d*) *Comparison with Allport's A/S test.* All of the males were given Allport's Ascendance–Submission Reaction Study at the time of retesting and scores were compared with self/dominant scores. When those cases with self/dominant disagreements of 5 or less and those above 5 were compared with those making positive scores ('ascendant') and those making negative scores ('submissive') a χ^2 value (corrected for continuity) of 4·29 was obtained which is significant at a 5% level of confidence:

	Self/dominant $\leqslant 5$	Self/dominant > 5
A/S positive	26	12
A/S negative	18	24

IV. COMMENT

Consistency of sorting was quite high for both groups, bearing in mind the unusual nature of the task. Recent work in which the form of the sorting question has been systematically altered suggests, however, that results are very sensitive to the precise formulation of the sorting question—showing that it would be desirable still further to reduce verbal intervention if this were possible. Another difficulty is that the recognizability of the faces seems to depend critically on their distance—and it seems necessary to specify this distance carefully. There were large individual differences among the psychiatric group in ease of perceiving the faces even at the standard distance and this effect seems worth following up since there was a strong suggestion that it was the 'organic' patients who found most difficulty in recognition.

In the interests of brevity right/left position effects were not controlled. The scoring method described holds good, however (both for 'agreement' and for 'consistence'), when each stimulus is compared with every other *twice* in the usual way—so that it would be possible to control this factor in more careful work.

Though the results of this inquiry have necessarily been given in group form the most useful way of employing the method is probably in the intensive study of the individual case. Thus the particular application discussed here—the 'Self Valuation Test' (Liggett, 1957*a*)—is intended to allow a study of the individual personality by the application not only of standard sorting questions but also of questions which the subject's own verbal responses to the portraits (or other sources of information) suggest as being particularly relevant. Suggestions for the systematic collection of these verbal responses and for other uses of this type of material are given elsewhere (Liggett, 1957*a*).

The technique for producing 'de-structured' stimuli is a relatively simple one and it should be possible to construct stimuli appropriate to many other quite different attitude problems. The method described may therefore prove to be a general one with other applications and uses. It is also hoped that the brief 'X-method' for paired comparisons may prove useful in other settings.

The writer would like to express his thanks to Prof. G. W. Allport for encouragement and advice and to Mr D. J. Newell, M.A., F.S.S., for statistical advice.

REFERENCES

ALLPORT, G. W. (1943). The ego in contemporary psychology. *Psychol. Rev.* **50**, 451–78.

CRONBACH, L. J. (1955). Processes affecting scores on 'Understanding of Others' and 'Assumed Similarity'. *Psychol. Bull.* **52**, 177–93.

FOULDS, G. A. (1948). The importance of an agreed subjective criterion in the interpretation of certain questionnaire and projective tests. *J. Personality*, **17**, 221–31.

HILGARD, E. R. (1949). Human motives and the concept of the self. *Amer. Psychologist*, **4**, 374–82.

KENDALL, M. G. (1955). *Rank Correlation Methods*. London: Griffin.

LIGGETT, J. (1957*a*). *The Self Valuation Test*. Newcastle upon Tyne: J. and P. Bealls.

LIGGETT, J. (1957*b*). *Translucent Comparison Sheets and Conversion Scales for the 'X-method' of Paired Comparisons*. Newcastle upon Tyne: J. and P. Bealls.

McLEOD, R. B. (1947). A phenomenological approach to social psychology. *Psychol. Rev.* **54**, 193–210.

(*Manuscript received* 10 *July* 1958)

THEORIES OF ANXIETY AND HYSTERIA—A REJOINDER TO HANS EYSENCK

By VERNON HAMILTON

Psychology Department, Springfield Hospital, London, S.W. 17

In his reply to my critique, Prof. Eysenck has attempted to reject the whole of my 'essay' on the basis of a discussion of only a few, and not always the most important, of my arguments. The points made by me have since received independent support by Storms & Sigal (1958, 1959). Further consideration of these and my paper, and Eysenck's replies (1959*a*, *b*), does not appear to strengthen Eysenck's case that introversion-extraversion is a viable continuum or base-line concept for subsequent testing of derived hypotheses, that his typological postulate has in fact been confirmed by the experimental findings presented by him, or that the various inhibitory concepts in the theory are in fact identical. Contrary to Eysenck's introductory remarks (1959*a*, p. 64), it would appear that he has not adequately and objectively concentrated on the main issues pertinent to his theories. I shall comment on Eysenck's points in the order in which they were made.

(1) When referring to dimensions and factors of introversion-extraversion obtained on the basis of 'factor-analytic predictions' from Jungian hypotheses, it is obvious that one is relating oneself not to Jung, who might not recognize these variants of his theory, but to Eysenck. Confusion would seem to be ruled out. While there is considerable objective and subjective evidence for Jung's general formulations, the Hildebrand study was, until this belated disclaimer, the major prop of the Eysenckian proof of introversion-extraversion, and the starting-point of further hypotheses. I have justified this conclusion by references on p. 49 of my original article. Eysenck does not refute these.

(2) My argument concerning the lack of unidimensionality of the major questionnaire measures of neuroticism and extraversion is based not on any factor loadings, although they reflect it too, but on the actual correlations between the measures. Storms & Sigal (1958), arguing essentially from similar premises, also show that the R scale significantly correlates with all the questionnaire measures of neuroticism so that an orthogonal solution of neuroticism and introversion-extraversion veils the undoubted correlation between the two variables. Eysenck now agrees (1959*b*, p. 66) that a significant correlation is present. To account for it by 'response set' as far at it applies to neurotics is disingenuous. The absence of unidimensionality becomes almost crucial, not only when selecting diagnostic criterion groups, but also when assessing the relative importance of the co-variants of the measures of reactive inhibition, i.e. neuroticism versus introversion-extraversion.

(3) Eysenck's argument at the top of p. 65 would seem to be a deduction from false premises. In a general population average I.Q.'s must be around 100, otherwise the concept of I.Q. becomes meaningless—the whole point of my argument was that apparently we do not know what constitutes average or non-neurotic scores on Guilford's R scale (viz. my table 2, 1959, p. 51). I considered that Hildebrand's and other normal control groups were not abnormally different from each other on extraversion because the group

means were separated by only 1 S.D., not 2 S.D.'s, and because variations in means between normal groups were less than obtained variations for hysteric groups (viz. *loc. cit.*). I therefore concluded (somewhat facetiously) that there was as much justification for rejecting Hildebrand's hysterics as there was for the rejection of the normal control group in this study.

(4) When discussing the doubtful selective validity of the Maudsley Personality Inventory, published evidence was not available to me. In the meantime, the critical assessment of the M.P.I. to which I was referring has appeared (Sigal, Star & Franks, 1958*a*) as well as a rebuttal of these criticisms (Eysenck, 1958). In their reply to this, Sigal, Star & Franks (1958*b*, p. 382) are able to reiterate their original conclusions based on the scores of groups of neurotic patients on the M.P.I. to the effect that '...either hysterics and dysthymics cannot be used as criterion groups in the measure of introversion-extraversion, or the E (extraversion) scale (of the M.P.I.) is not a good measure of introversion-extraversion, or both statements are true'.

(5) Eysenck's theory demands not only that dysthymics should have lower scores than hysterics or psychopaths on a *good* measure of extraversion (Eysenck, 1959*b*, p. 65), but that the relevant neurotic scores should be *significantly* different from each other and from control groups. Sigal *et al.* (1958*a*, p. 145) did not find what Eysenck claims they found (Eysenck, 1959*b*, p. 65, footnote). They found *no* significant differences on the E scale of the M.P.I. between the following groups: dysthymic *v.* hysteric/psychopath, normal *v.* hysteric/psychopath, normal *v.* hysteric, normal *v.* psychopath, dysthymic *v.* hysteric. In other words, the M.P.I. is an inadequate device for obtaining *criterion groups*. And that is the crux of my argument: the M.P.I. lacks proven external validity and one therefore cannot be sure whether selection of neurotic introverts and extraverts by the scale will yield such groups. Consequently any differences between such selected groups on other parameters cannot be related to introversion-extraversion or to dysthymia-hysteria.

(6) On pp. 66–67 of his reply, Eysenck suggests that I have misunderstood the function of a postulate in a hypothetico-deductive system, and that I have ignored important evidence in support of his typological postulate apart from the experiments on conditionability and kinaesthetic figural after-effect. With regard to the evidence which lack of space prevented me from considering, I can do no better than refer the reader to the papers by Storms & Sigal (1958, 1959). In their discussion of Eysenck's theories, these authors produce considerable evidence to show that the deductions which Eysenck makes with regard to sedation thresholds, inert versus over-active behaviour, dark adaptation, apparent movement, rotation effects and CFF, are not clearly supported by the evidence cited by Eysenck, and that there is in fact considerable doubt that they *can* be deduced from the theory (1958, p. 243).

With regard to the validity of methods of constructing postulates or principles, I think Eysenck is on rather shaky ground generally if he uses the Hullian system as a paradigm. Hull has been criticized on the grounds that his principles have variously contained simplifying assumptions, conventions concerned with quantification and generalizations from experimental data, apart from the basic formulations in terms of hypothetical variables. In the introduction to *Essentials of Behaviour*, Hull (1951) writes: 'The primary laws of the present behaviour system are all empirical in nature, i.e. they have all originated in one way or another from observation and experiment.' It goes without saying that a postulate must go beyond these. The question is how far, and with how much previous observation and experiment.

While many of the assumptions underlying Hull's postulates are only indirectly testable, some of Eysenck's may clearly be tested directly and with advantage to the theory. On the other hand, I brought evidence which is contrary to assumptions made in Eysenck's postulates, viz. the lack of evidence for the identity of negative induction, reactive inhibition and satiation (p. 58 of my paper). The whole argument would seem to amount to this: that there are postulates and postulates, all of them on a continuum from those derived by free-association to those having maximal empirical support. I am saying that Eysenck's postulates do not belong to the latter category.

During a recent symposium on inhibition (*Bull. Brit. Psychol. Soc.*, May 1959), considerable disquiet over the adequancy of the Hullian system was expressed. Attention was also drawn to present tendencies in work on personality within a framework of Hullian formulae, to ignore the 'K' factor, thus considering excitatory and inhibitory states of the organism separately and without reference to motivation and incentive conditions. This must be considered a serious omission in work with *any* group of subjects.

(7) It is true that in Franks's conditioning experiment subjects who blinked in response to three pre-experimental presentations of the tone were excluded from the trial. My statement that at the beginning of the conditioning experiment there was already a large difference in strength of C.R. acquisition between hysterics and dysthymics was therefore ambiguous. I wished to suggest that the substantial difference in eye-blink conditioning between hysterics and dysthymics on the first experimental trial may have been due to inadequate pre-experimental testing with the C.S. (Storms & Sigal, 1958, p. 235, suggest that '...tone alone trials previous to conditioning should have consisted of eighteen presentations in order to provide comparability to the test series scores'), and/or that differences in *experienced* stimulus intensity (UCS) would determine by themselves differences in C.R. acquisition at any one point of a series of experimental trials and especially in the early trials. In terms of this variable, there is no evidence that Franks's groups were equated.

(8) On the relationship between conditionability and degree of neurosis, Eysenck has not met my argument. Of course, there would be no significant difference in number of C.R.'s produced between neurotics and controls in Francks's study if the latter lie intermediate to the two neurotic subgroups whose high and low C.R.'s would average out. That evidence is irrelevant when considering the degree of neurosis or neuroticism in the individual subjects making up the two neurotic groups. We simply need to know whether the hysterics were less neurotic than the dysthymics. I think they were, but my only evidence is in terms of scores on the R scale which, it will be remembered, is inversely and significantly related to neuroticism.

Recent evidence (Bartholomew, Franks & Marley, 1958) shows nearly zero correlation between neuroticism and conditionability but also a non-significant correlation between extraversion and conditionability, while Gwynne Jones reported (B.P.S. Annual Conference 1959, Symposium on Inhibition) that recent work at the Maudsley Hospital obtained a relationship between salivary, spatial, verbal condition and neuroticism, not introversion-extraversion, and that a general factor of conditionability was not found.

(9) It is not true to say, as Eysenck does (p. 68 of his reply), that his 'original demonstration of the relationship between kinaesthetic figural after-effect and extraversion'

has been verified by 'several investigations'. Reference to Eysenck's summary of studies on the relationship between kinaesthetic figural after-effect and extraversion-introversion (1957, p. 155ff.) will show that he reports only those studies—his own and Nicholl's—which I have already critically considered. Eysenck has also agreed now that Broadbent (1958) has shown that figural after-effect may be an experimental 'artifact'. My original criticism thus remains: there is no empirical evidence to link reactive inhibition, satiation or kinaesthetic figural after-effect (which might have been obtained from a correlational study), and there is insufficient evidence to link these variables with the typologies. While there is undoubtedly a *similarity* between the first two concepts, Eysenck postulated quite a different degree of relationship: *identity*. This, as Eysenck says, is difficult to demonstrate. Since he has not demonstrated it, and since his explanation of this failure in terms of involuntary dissipation of inhibition remains at the observational level, I suggest that this aspect of his postulate may need revision.

(10) With regard to the relationship between reminiscence effects and introversion-extraversion, I apologize to Eysenck for misreading the footnote on p. 254 of his recent book (Eysenck, 1957). No work appears to have been published or even carried out with neurotic subjects and my remarks conveyed a false impression. It is pertinent to ask, however, why this obvious deduction from theory was not submitted to experimental testing. I shall hazard a guess that Eysenck will draw attention here to the dissipation of inhibitory potential which may occur with extraverts, and which would make it difficult for reminiscence effects to appear. If this is true, as it well may be—and as I said before (1959, p. 60) we have no empirical evidence for it—then the findings supporting the theory with respect to *non-neurotic* extraverts and introverts are unacceptable for the same reason. This alternative explanation, however, if correct, would also allow one to invoke low drive as a factor in the production of marked reminiscence effects. Furthermore, in a re-analysis, using different criteria, of Eysenck's own reminiscence data (Star, 1957, pp. 135ff., the details of which cannot be given here), reminiscence correlated significantly with neuroticism and not with extraversion. One would thus be inclined to agree with this investigator who writes at the end of a comprehensive investigation into the relationship of reminiscence and introversion-extraversion that '...from the experiments reported we are not able to arrive at a complete understanding of the way in which the extraversion and neuroticism traits relate to reactive inhibition' (Star, 1957, p. 180).

REFERENCES

BARTHOLOMEW, A. A., FRANKS, C. M. & MARLEY, E. (1958). Susceptibility to methylpentynol: eyelid conditioning and P.G.R. response. *J. Ment. Sci.* **104**, 1167–73.

BROADBENT, D. E. (1958). Extraversion and the kinaesthetic figural after-effect. *Bull. Brit. Psychol. Soc.* no. 35.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1958). Hysterics and dysthymics as criterion groups in the study of introversion-extraversion: a reply. *J. Abnorm. Soc. Psychol.* **57**, 250–2.

EYSENCK, H. J. (1959*a*). Scientific Methodology and '*The Dynamics of Anxiety and Hysteria*'. *Brit. J. Med. Psychol.* **32**, 56–63.

EYSENCK, H. J. (1959*b*). Anxiety and hysteria: a reply to Vernon Hamilton. *Brit. J. Psychol.* **50**, 64–9.

HAMILTON, V. (1959). Eysenck's theories of anxiety and hysteria: a methodological critique. *Brit. J. Psychol.* **50**, 48–63.

HULL, C. L. (1951). *Essentials of Behaviour.* New Haven: Yale University Press.

SIGAL, J. J., STAR, K. & FRANKS, C. M. (1958*a*). Hysterics and dysthymics as criterion groups in the study of introversion-extraversion. *J. Abnorm. Soc. Psychol.* **57**, 143–8.

SIGAL, J. J., STAR, H. & FRANKS, C. M. (1958*b*). Hysterics and dysthymics as criterion groups in the measure of introversion-extraversion: a rejoinder to Eysenck's reply. *J. Abnorm. Soc. Psychol.* **57**, 381–2.

STAR, K. H. (1957). An experimental study of 'reactive inhibition' and its relation to certain personality traits. Unpublished Doctor's dissertation. University of London Library.

STORMS, L. H. & SIGAL, J. J. (1958). Eysenck's personality theory with special reference to '*The Dynamics of Anxiety and Hysteria*'. *Brit. J. Med. Psychol.* **31**, 228–46.

STORMS, L. H. & SIGAL, J. J. (1959). Misconceptions in 'Scientific Methodology and *The Dynamics of Anxiety and Hysteria*'. *Brit. J. Med. Psychol.* **32**, 64–7.

(*Manuscript received* 23 *April* 1959)

CORRIGENDUM

In 'Attention to visual displays during a vigilance task. II' by C. H. Baker, Vol. 50, page 31, the legend under Figure 1 should read as follows. 'Showing the percentage of pips missed of the total number of pips *missed*, as a function of time, for six experiments. Mackworth's data (1944) are plotted for comparison purposes and show, for each half hour, the percentage of signals missed of the total number *presented*.'

PUBLICATIONS RECENTLY RECEIVED

A Psychological Study of Typography. By Sir Cyril Burt. Cambridge University Press. 1959. Pp. xix + 68. 15*s*.

The greater part of the text of this book appeared previously as an article under the same title in the *British Journal of Statistical Psychology* (by C. Burt, W. F. Cooper and J. L. Martin, 1955, **50**, 29–57). Since, however, the material is of considerable interest—to the general reader as well as the psychologist—its appearance as a separate publication is very welcome. The original form, that of a research report, has been largely preserved, and Sir Cyril is careful to disclaim generality for his results beyond the scope of the inquiry, which was primarily concerned with children's readers and scientific periodicals, so far as questions of legibility were concerned. On the other hand, the parts of the work dealing with preferences carry wider implications. Sir Cyril rightly calls attention to the changes in typographical preferences revealed by a comparison of his recent results with those of an earlier study, published in 1933. It may be of interest, too, to note that an order of preference for eight type faces (all included in Burt's study) recently published in the Edinburgh University *Gazette* shows a correlation of only +0·24 with Sir Cyril's, and that this figure is entirely accounted for by the low popularity, in both groups, of Walbaum. Such discrepancies, however, do not in any way invalidate Sir Cyril's main conclusion, which is that 'printed matter seems more legible, and reading becomes more accurate and quick, when the material is set in a type which the reader, perhaps without realizing it, finds aesthetically pleasing'.

This is particularly interesting inasmuch as few readers—even among the sophisticated—are 'type-conscious'. As Stanley Morison (who is Typographical Adviser to the Cambridge University Press and designer of the immensely successful Times New Roman type face) writes in his illuminating Introduction: 'That so many kinds of type look so much the same to so many people is a tribute to the skill and sense of the successors of the masons and scribes who centuries ago decided the form of the majuscules and minuscules.' Mr Morison believes that 'this lack of typographical consciousness and knowledge is welcome'—on the grounds that the function of the printer's art is not primarily to delight the eye, but to afford the reader 'the swiftest access to the author's thought'. If that is so, it is sad that increasingly often nowadays, even in the fields of serious publication, the achievements of this same skill and sense are jettisoned, for reasons of economy or convenience, in favour of mimeographed typescript or similar make-shifts. Itself a notable example of fine book-production, the *Psychological Study of Typography* may perhaps help to arrest this trend.

B. SEMEONOFF

Television and the Child: An Empirical Study of the Effect of Television on the Young. By H. T. Himmelweit, A. N. Oppenheim and P. Vince. London: Oxford University Press (for the Nuffield Foundation). 1958. Pp. xix + 522. 42*s*.

The title of this book is somewhat of a misnomer, since it suggests that the results of the inquiry described have a greater generality than they in fact possess—or indeed than the authors claim for them. Children in two age groups only—10–11 and 13–14 years—were studied; and these by questionnaire methods, not individually. Thus it was not possible to make any real study of the effects of television on younger children. Inevitably, owing to the limited period of time allowed for the whole inquiry, no very long-term results could be obtained. However, the study of the Norwich children, before and after television became available there, gave a valuable check on the remainder of the results by showing that children of families which bought television sets after television was available differed beforehand from those in families which did not later buy television sets.

However, the imperfections of the study, many of them unavoidable, should not deter us from praising it highly, especially from the methodological point of view. The technique of matching each viewer in the samples with a non-viewer of the same age, sex, intelligence and socio-economic group went far to ensure that the differences observed between viewers and controls were genuine. Clearly these were more often than not due, not to the watching of television itself, but to the fact that the children who view, especially for long periods (the television 'addicts'), were different in nature from those who did not; the former tended to be duller, more passive and more 'amusement oriented', and sometimes more withdrawn and unsociable. It is possible that excessive viewing exaggerated these tendencies, but in general television 'addiction' was a symptom rather than a cause.

The main data utilized in the inquiry were obtained by means of a series of questionnaires, teachers' ratings of personality and school performance, tests of the children's information, diaries and programme recall lists. The diaries were kept by the children of their activities during each leisure hour of the day over a period of one week. Though we cannot be certain that all relevant items were accurately remembered and recorded, every precaution was taken to ensure this. Unfortunately, the investigators did not have time to work out the results from all the diaries. The questionnaires were most skilfully drawn up, so that some of the more important problems were approached from several angles. For instance, the incidence of acts of violence in televised plays and films, and the children's reactions to them, were covered by analyses of the type of violence appearing in the programmes most often viewed; and by questions as to whether the children had ever been frightened by such incidents, whether they liked or disliked them, whether there had been any incidents which did frighten them, whether there were any incidents which were bad for children to see, what they would do to make a programme exciting or frightening. Again, in order to avoid undue emphasis on reactions to television, similar questions were asked about the cinema, radio, reading matter, etc.

Several additional studies were made, including those of the opinions of parents and teachers as to the effects of television on children. The answers to these, however, seemed to be too closely related to the parents' and teachers' own personal reactions to television to afford much reliable evidence. Some interviewing of parents and teachers, and observations of children's behaviour while viewing, were also carried out. But the numbers were necessarily very small; and therefore the information obtained, though doubtless helpful to the investigators, could not justifiably be generalized. Indeed, it would have been better to omit it. Finally, an interesting experiment was carried out to indicate the effects of a children's television programme about the Ashmolean Museum at Oxford on the subsequent visits of the children to that Museum; the effects appeared to be nil!

Perhaps the most important outcome of the inquiry is that it presents a broad picture, sociological rather than psychological, of the general place of television in the lives of children of the types studied. In fact, the viewing of television seems to take its place as one form of 'admass' entertainment. Most children learn little from it; but neither does it do them any particular harm. It does not frighten them much, nor make them more aggressive, nor take up an undue amount of their time. Over a long period of viewing, the values presented by certain popular programmes—the stress on self-confidence, toughness and violence, and lack of importance attached to the gentler virtues—may possibly reinforce the tendency to emphasize similar values apparent in other types of mass entertainment. The more intelligent children, though they are more critical and view less than the less intelligent, may suffer from a general flattening down of educational and cultural interests and activities. It is notable that the grammar-school viewers were assessed by their teachers as significantly poorer than the controls in several of the main school subjects.

But what are needed now, as the authors themselves point out, are long-term studies of individual children's reactions to television, especially those of young children. Do they form habits over the years of greater aggressiveness or timidity, of less cultural interest and greater amusement orientation? On the other hand, if they were shown more and better instructional programmes, would their interests be broadened? Or will the effects depend almost entirely on the children's personalities and home backgrounds, temperamentally unbalanced children being more affected than others? In sum, is viewing television just one feature of children's ordinary lives nowadays, not significant in itself but merely a part of the general pattern of modern civilized life?

We must congratulate the authors of this book in that they have obtained, in the all too brief period allowed them, a wealth of valuable information on the effects of television on certain types of children, and have presented it with admirable clarity and in a manner which will interest both the ordinary reader and the professional sociologist and psychologist. Perhaps the inferences they have made sometimes go beyond the evidence. Further and more intensive studies are needed to confirm or invalidate these inferences.

M. D. VERNON

Psychiatry in the British Army in the Second World War. By ROBERT H. AHRENFELDT. London: Routledge and Kegan Paul. 1958. Pp. xvi+312. 35*s*.

At last, the changing mental climate in regard to the field of psychological medicine has made possible and acceptable this very excellent historical conspectus of the applications of psychological medicine (and to a lesser extent of psychology) in the Second World War.

Dr Ahrenfeldt is thrice lucky. Writing six years after the official medical historians and thirteen after the end of the war; having wide and long experience as an army psychiatrist as well as post-war access, as

Deputy Assistant Director of Army Psychiatry, to all the documentation and records; availing himself thoroughly of the considered and weighed present opinions of all those best competent to evaluate their experience and the available material evidence. He has used his luck as wisely as he has used his material. No one who has anything to do with the maintenance of morale, 'moral and physical guts' and resistivity to the disintegrating effects of the very excessive stresses on Soma and Psyche ('human beings' would be a less pretentious and more comprehensive term) in warrior or worker can afford to dispense with this book.

Dr Ahrenfeldt makes it clear—though, necessarily perhaps, he is ultra cautious and conservative in opinions expressed (which are few)—that the Army made a particularly good job of those of the subnormal 10% who got into the Army. Instead of constituting a grievous burden of semi-demented or 'bloody-minded' men, they became, for the most part, a cheery lot, stimulating of morale when one met them by their obvious enjoyment of the very useful work they knew so well how to do.

Personnel Selection below officer level is very adequately handled. Officer Selection is especially well handled; as I can vouch from personal experience. Inevitably, because of the theme of the book, dates and history loom large. Actual theory, practice, principles and the relevance of all these to present Service and Civilian needs are not considered; as indeed it would be unwise of the author to consider here. One hopes it will refer those committed to contemporary projects to new and important work that has not been adequately appreciated, understood or made use of.

Hospital treatment of broken-down soldiers probably made least progress in war time; though considerable man-power saving economies resulted. There was valuable experimentation in group-therapy; especially valuable because it aimed, where possible, to return men quickly to jobs where they were of use to the Army. Apart from those treated, within days, at Forward Centres, they rarely returned to combatant service.

The handling of disciplinary problems was always incredibly difficult. Inevitably it must have been difficult for warrior officers to accept the advice of non-warrior doctors; especially in a field of medicine so unexplored as yet. From my own relatively small experience in such matters, I must confess what amazed me was the tendency, after a time when the doctor proved his worth, of some commanding officers to send all, or nearly all, their difficult disciplinary cases to the 'trick-cyclist'. Was it a feeling of uncertainty, or a desire to get the 'psych' to do his scavenging for him? Most discharged psychopaths earned many a drink boasting how they fooled that innocent of innocents, the 'Area Psych'. Were their C.O.'s to be interrogated, the story was likely to be different. Naturally 'psych's' and C.O.'s got much fun from retailing such stories to each other.

I must admit what interested me most was the brilliant, and entirely original, and—to my mind—almost the most valuable work of all, where the ultimate welfare of the post-war civilian community was concerned, the rehabilitation and civil resettlement of expatriated prisoners of war. I have no doubt that an application of some of its principles to existing institutions dealing with employment, and possibly to patients discharged from the ordinary hospitals, could help men, women and children of the community to adjust in the interstices between one job or school or training institution to another: even from one age period to another, as from school to job, job to retirement or old age to advanced old age, where economic and emotional demands change abruptly.

After such a good job of work, the author might have relaxed from pseudo-scientific 'scepticism' which regards an expression of strong, but not inflexible, belief in his conclusions as a sign of bad taste and flabby thinking. There is no evidence that the scientific minds of highest quality have 'emoted' flabbily. Galileo 'emoted' quite strongly. I feel the author, after so much contact with his subject, must feel the Army did a marvellous job of work. So marvellous, that he might have ventured to suggest the next lines of application and research, for the community in general, whose use of its leaders, its best brains, its inferior brains, its damaged personalities (some damaged for the lack of precisely such techniques), is still behind the optimum practices of the Army in wartime, and certainly behind what can and should be done.

HENRY HARRIS

Psychological Stress: Psychoanalytic and Behavioural Studies of Surgical Patients. By Irving L. Janis. New York: Wiley; London: Chapman and Hall. 1958. Pp. xiv + 439. 56*s*.

This is probably a further milestone towards a plausible partnership between psychodynamics and the main stream of experimental practice in psychology. While at first seeming rather long, the book is so well organized that the author persuades us it could not have been curtailed. The more imaginative psychologist can be assured of stimulation to some long thoughts—the less said of the likely response of his more rigid brethren the better! Since the style is comparatively free of jargon, it's 'a possible' for surgeons, who should begin with chapters 17 and 26.

The book falls naturally into two parts. The first part (chapters 3 to 17) utilizes psychoanalytic material from a patient fortuitously undergoing surgery. Chapters with almost verbatim accounts of pre- and post-operative therapeutic sessions alternate with others formulating internal mediating processes under fourteen specific hypotheses.

These hypotheses employ the central concept of regression. Exposure to body damage unconsciously revives threats of punishment in childhood. The primary defences outlined include: (i) diminishing own aggression for fear of retaliation, (ii) repression of impending surgery together with displacement of responsibility for any residual bodily anxiety, (iii) dwelling exclusively on unrealistic compensatory gains, (iv) transference onto 'authorities' (surgeons, etc.), (v) post-operative euphoria or depression and resentment depends on whether more or less suffering is experienced than anticipated; this in turn is coupled with the perceived identity of the 'salient present' with relevant infantile experiences, the duration of stress, self-punishment needs and childish memories of irrational punishment. These defences naturally break down under the impinging realistic details of approaching surgery.

The second part (chapters 18 to 26) consists of data derived: (i) from pre- and post-operative interviews (together with behavioural ratings by medical staff) of 22 patients, (ii) retrospective questionnaire data from 200 male university students. Three further hypotheses are advanced to the effect that patients showing (*a*) extensive, (*b*) little or no, pre-operative fear develop severer emotional disturbances during or after stress than (*c*) patients displaying a moderate amount of pre-operative fear. The first group tend to develop post-operative anxiety and depression—the second group tend to display post-operative resentment to authority. The pre-operative fear of the third group seems to constitute an effective 'working through' to ensure adaptation in the post-operative phase.

The book probably lacks modesty in its title 'stress'. As the author admits, stress refers specifically to surgery, in which (i) the subject has to be passive, (ii) has to rely on an authority figure (surgeon), both of which tend to congruency with regression and infancy. One also wonders to what extent one psychoanalytic case can suggest the whole gamut of defences operative. However, the book is an effective incitement to further research, an amusing 'stick' with which to provoke surgeons to a further awareness of the psychological—and, perhaps, to make psychologists aware of the surgical? ARNOLD BURSILL

The Older Worker in Industry. By G. H. CROOK and M. HEINSTEIN. Berkeley: University of California Institute of Industrial Relations. 1958. Pp. vi + 143. $2.00.

This book seems to have been published prematurely. It is based on a study of attitudes and opinions expressed during interviews with a sample of '846 industrial workers in the San Francisco and Los Angeles areas, conducted in 1952–1954'. The interviews are supplemented by data from supervisors about the interviewees' work performance and by factual data from the records of their employing firms. Unfortunately, this last type of datum was not fully analysed and related to the rest. As a result, the book presents a series of rather 'flat' lists of the percentages of men and women over and under fifty years of age who give particular answers to items in a questionnaire. The information is essentially subjective in nature and, although it is potentially interesting to social workers and to those concerned with reactions to retirement, it is very difficult to be sure what reliance can be placed on it.

The study was made with the support of the Rockefeller Foundation as part of an attempt to answer the question: 'At what age and under what conditions should Society retire its aged members?' The method adopted was, unfortunately, not the best that could have been used in this attempt. What is required would seem to be a close study of changes in attitude to retirement as people approach retiring age rather than a comparison of two groups, one of which contains substantial numbers as young as the twenties and thirties. The shortcomings of their data have given the authors the task of 'making bricks without straw' in their discussion. This is, in consequence, somewhat long-winded and excessively repetitious. It does, however, show a penetration and insight worthy of much better data. We may hope that the authors will have the opportunity of working out their results fully and presenting a complete report in due course, although to do so would admittedly need a great deal of difficult work. Perhaps the moral to be drawn from this book is the one which will be familiar to anyone who has worked in the field of ageing, namely that questionnaire survey studies are an easy method of getting a certain amount of superficial information but are an extremely laborious and often expensive means of obtaining any real understanding.

A. T. WELFORD

Measuring Security in Personal Adjustment. By M. D. and L. H. AINSWORTH. Toronto University Press; London: Oxford University Press. 1958. Pp. xi+98. 24*s*.

This book reports on the construction and use of a group of four paper-and-pencil tests of personal adjustment based on the theories of Dr W. E. Blatz, Director of the Institute of Child Study, University of Toronto. In his theories, the concept of security plays a leading part. The first researches towards the present group of tests were published in 1940, and since then the work has progressed in the hands of the authors of this book and various co-workers.

It is intended to develop tests of security and insecurity in five directions of personal adjustment. These are: (*a*) parent-child relationships; (*b*) relationships outside the primary family; (*c*) adjustments to work or school; (*d*) leisure-time activities; (*e*) religion and philosophy of life, and attitudes to health, disease and death. In the present book, tests in four of these directions are given, omitting adjustments to work or school. The tests, which are quoted at the end of the book, involve seven pages of questions, and, as the authors say, they must continue to undergo revision and study. The present forms of them are suitable for young adults of college age. The book is a very interesting study in the steps, stages and processes of test construction and validation, and the tests themselves should be useful to practical psychologists.

R. W. PICKFORD

The Integration of Behavior. Vol. III: *The Reintegrative Process in a Psychoanalytic Treatment.* By THOMAS M. FRENCH. University of Chicago Press; London: Cambridge University Press. 1959. Pp. xii+484. 75*s*.

This book is the third volume of a series of five called *The Integration of Behavior*, of which the first deals with *Basic Postulates* (1952) and the second with *The Integrative Process in Dreams*, while the remaining two will deal with *Factor Analysis of Behavior* (not factorial analysis) and *Biological Foundations of Behavior*.

While the present volume is written with admirable clarity, and no effort has been spared to explain every detail of the steps in the treatment of the patient discussed, it is almost too long and repetitive. Some readers would prefer to pay less than £3. 15*s*. for a more concise account, and, when it is realized that only a single case is presented, the reviewer wonders wether something more efficient might not be done from the point of view of the reader. On the other hand, it is clear that even a single case can never be presented absolutely fully, however long a report is given of it, and in spite of reading the whole book, still we have not met the patient. However, the essential ideas are contained almost entirely within Section I: 'Theoretical Analysis of Psycho-analytic Therapy'; and also in Section III: 'Recurring Sequences and Therapeutic Process'. The remainder of the volume may be said to elaborate and illustrate the points brought out with great lucidity in these sections, especially the first—only 44 pages.

There is no doubt that many non-psychoanalytic readers will find the volume most stimulating and helpful, and the very thoroughness of the explication will make them feel more confidence in the theories and methods. Analytic-minded readers and experts will be deeply interested in considering the implications of a case of asthma such as that described, from a therapeutic point of view, and of the approach made to the patient, together with the details of method and theory deduced from the case and illustrated by it.

The central principles illustrated in the volume are: (1) that neurosis results from a hold-up of the normal development of rational goal-seeking activities under the influence of intense conflicts; and (2) that in the therapeutic situation this developmental or learning process may be re-activated and, under suitable conditions, carried to completion. In a sense these points are obvious, but the author's great gift for insightful handling and interpretation of the patient's problems and how they were met in the analysis, fills his volume with an ever new and refreshing interest.

R. W. PICKFORD

Family Influences and Psychosomatic Illness: An Inquiry into the Social and Psychological Background of Duodenal Ulcer. By E. M. GOLDBERG. London: Tavistock. 1958. Pp. xii + 308. 38*s*.

The clinical approach to psychosomatic disorders seems to have reached its limit as a research technique. Other specialities and the basic sciences are taking over and are currently following up clinically derived hypotheses and hunches in the laboratory and in the field. Miss Goldberg, an experienced social inquirer, here reports the exploratory stages of a socio-psychological investigation of family relationships, personality development and duodenal ulcer. In collaboration with a psychiatrist and clinical psychologist she has undertaken what might be called a 'middle scale' study (too big to be called clinical and too small to be called statistical!) of a group of young males diagnosed as suffering from duodenal ulcer and a randomly selected group of controls matched for age. The book states clearly that the aim of the investigation is to produce hypotheses and not 'definite findings of general validity'. In this respect the dimensions of the study allow the use of a technique of inquiry which involves isolating variables without destroying the 'inter-relatedness of specific aspects of behaviour and attitudes in actual people and families'.

What emerges is the tentative hypothesis that duodenal ulcer in young men appears when they attempt to leave a carefully regulated home and family in which mother is the home-maker and authority. It is the son's conflict between dependence upon this powerful mother and the demands of adult life which is basic in duodenal ulcer patients. Such an hypothesis has been stated before by clinicians. The fact that it is derived in this context from a study of family background and presented in such a disciplined manner gives it a new significance.

It is difficult to decide whether such aetiological factors are specific to duodenal ulcer patients, to all psychosomatic conditions, or could be due to something more general still—such as neuroticism. The results of the psychological tests reported in an appendix certainly indicate a high degree of neuroticism in the group of duodenal ulcer patients. It would seem that some refinement of the hypothesis and the use of more precise test procedures will be necessary for the next stage of the investigation. Readers of this book will have no doubt that the necessary skill and competence are being used to tackle the deductive phase of this study.

The value of this book will be enhanced by future reports from this research team. For the moment, anyone concerned with problems of case study method should read it. R. M. MOWBRAY

Die Prozent-Quotient-Methode. By GERHARD CLOSTERMANN. Münster: Aschendorff. 1958. Pp. 91. No price given.

This short monograph deals with a new method of calculating coefficients of contingency. This new coefficient is similar to the original Yule formula except that it uses the sum instead of the product of the opposing quadrants of a fourfold table. The author gives tables for reading off the coefficient which should make its determination very easy indeed. The coefficient does not attempt to estimate product moment correlations but must be interpreted in its own right. Its values, on the whole, tend to be very similar to those of the Yule coefficient except when correlations are very high, when the latter tends to be higher. Both the new coefficient and Yule's coefficient tend to be markedly higher than tetrachoric coefficients or product moment coefficients.

The author gives many examples to illustrate the use of his coefficients; to many readers, these will be even more interesting than the coefficient itself. Thus in one table, the author demonstrates the strong correlation between hair colour and tendency to develop scarlet fever. In another example he gives correlations between the body build of the husband and that of the wife, showing that leptosomatic men do not show significant preferences, while athletic men prefer athletic women quite strongly; pyknic men show slight preferences for pyknic women. In another table, the author demonstrates the existence of a very weak relationship between juvenile delinquency and quarrelling between parents. The short monograph is well worth having, although the new method will probably be used in the main for trial runs and for obtaining a quick survey of results rather than for their more detailed treatment. H. J. EYSENCK

Stuttering. By C. F. Diehl. Illinois: Thomas; Oxford: Blackwell. 1958. Pp. xv+314. 73*s.* 6*d.*

This is a compendium. It is a well-nigh perfect demonstration of a low-level but psychologically literate skill. It is much more beautifully produced than *Psychological Abstracts*. It is also much more expensive. It is of no great value to anyone in this country. It is, indeed, hard to see who would find it useful except Prof. Diehl's 'average college student' who 'may select the general topic of "stuttering" for a research paper'. It will enable him to delude himself that primary sources are not important.

The compendium contains 193 abstracts from the mass of articles on stuttering (stammering). They are nearly all American, and are organized in five 'units' as follows: History (10 pp.); Symptomatology (57 pp.); Etiology–Physiological (88 pp.); Etiology–Psychological (63 pp.); Therapy (31 pp.). There are 364 additional references by title only. In themselves, the abstracts are clearly and systematically presented. Some are worthless in the light of later work. Others are worthless in themselves, e.g. 'the data...seemed to indicate that the saliva of stutterers is not significantly different from that of normal speakers, and that pH change should prove of little value in diagnosing stuttering cases'. The compiler, to be sure, indicates that the publications were selected as being 'representative of certain areas of research and theory', and not for their excellence.

G. SETH

Instinctive Behavior; The Development of a Modern Concept. Trans. and ed. by C. H. Schiller. London: Methuen. 1959. Pp. xix+328. 45*s.*

This book consists of eight articles by ethologists, none of which have appeared in English before. Six of the articles are translations of papers by Lorenz and Tinbergen, and they include Lorenz and Tinbergen on the egg-rolling behaviour of the grey-lag goose and Tinbergen and Kuenen on the gaping behaviour of young thrushes. There are four theoretical articles by Lorenz himself ranging from 1934 to 1952. Of the remaining two articles, one is a semi-popular account by von Uexküll of the sensory world of animals, and the other is a hitherto unpublished account of some experiments on insightful behaviour in chimpanzees by Paul Schiller. Since the Lorenz and Tinbergen articles are much the most important, attention will be concentrated on them in the remainder of this review.

The book provides an excellent history of the development of the ideas of the comparative ethologists. Through its pages we can trace the development of such ideas as the innate releasing mechanism, the consummatory act, the lack of plasticity of instinctive behaviour, the distinction between consummatory act and taxis, etc. Seeing these ideas set out in the context of the observations and experiments which led to them may make psychologists more sympathetic towards the ethologist's explanations of instinct. There is no tendency in these original articles to overgeneralize or to oversimplify: both tendencies appear in the published books of ethologists. Psychologists have tended to ignore the actual experimental results of the ethologists, possibly because many psychologists, through language difficulties, may not have read the original experimental papers and therefore have not appreciated how carefully controlled some of the experiments were. Later criticisms are often forestalled by Lorenz himself; e.g. he mentions that the hydraulic model is unable to account satisfactorily for the massive discharge of action specific energy in vacuum activity and that he may sometimes have mistaken a raising of threshold for a consummatory act due to the habituation of a specific part of the releasing mechanism for a raising of threshold due to the discharge of action specific energy. Many of the warnings in the book go unheeded by psychologists: for instance Lorenz stresses over and over again that instinctive behaviour is likely to degenerate in a domestic strain of animals, yet this has not prevented American psychologists from attempting to test for such things as imprinting and releasing mechanisms in domestic animals.

There are two features of the articles which may strike a psychologist as undesirable. First, Lorenz is given, at times, to unnecessary metaphysical speculations: he is anxious to show that animals are not machines, and uses words like 'subjective' and 'voluntary' of animals. However, this is a minor blemish and scarcely affects the main trend of his thinking. Secondly, the ethologists do not seem to have taken their own theories seriously enough: having set up an explanatory model, they do not try to draw prediction from it and then attempt to break it experimentally. For example, it is maintained that the consummatory act cannot be modified, but as far as I am aware no attempt has ever been made to modify the consummatory response by training. Again an alternative to the turning off of a specific drive by the release of energy is the turning off of a drive by feed-back from a stimulus and this is something which is known to occur in mammalian behaviour. Yet the ethologists have not attempted to test this alternative explanation by systematically eliminating the stimuli which usually follow or accompany a given consummatory act.

The articles are excellently translated by Claire Schiller, and should be widely read by psychologists interested in animal behaviour. If this book helps to promote a better understanding by psychologists of the ideas and approach of the comparative ethologists it will have served a very useful purpose.

N. S. SUTHERLAND

An Introduction to the Behaviour of Invertebrates. By J. D. CARTHY. London: Allen and Unwin. 1958. Pp. xi+380. 45*s.*

Although a vast amount of information about the senses of invertebrates has accumulated over the last sixty years, it has not previously been brought together in an accessible English text-book. Dr Carthy has successfully achieved this almost encyclopaedic task: his book will be a useful source of references for many years to come. Like all directories, it tends to be comprehensive at the cost of selectivity, and lacks any unifying thesis. Indeed, the multitude and diversity of the facts presented almost precludes cohesiveness.

Dr Carthy's approach is that of classical zoology—correlating structure with function. He stresses the necessity of tailoring experiments to fit the natural life of the animal, and the need to relate interpretation to normal functional significance: psychologists could well benefit from this reminder. In other respects, however, this attitude may become restrictive. Although in sensation and perception studies, correlating behaviour with the structure of sense organs is an important preliminary, the precise analysis of stimulus parameters and modes of function has to be made electrophysiologically. In this book, the physiological aspects of sensory activity are largely ignored, which is unfortunate. A description of the E.R.G. and of On-Off Fibres, for example, particularly if allied to Hassenstein's elegant (but here unmentioned) model would make sense of the theories of form perception by flicker. Similarly, Dr Carthy adopts (and in fact mis-states, p. 121) the mosaic theory of acuity, which was based on anatomical foundations, and which has been untenable for many years in the light of physiological results: the fact that acuity sometimes equals the ommatidal angle sets rather than answers the problem of the underlying mechanism.

The book also, and deliberately, omits any account of the details and organization of complex innate or learned behaviour. This is also unfortunate, for it is in this field, and in neurophysiology, that the most vigorous, exciting and important work has been done during the last ten years. It was therefore misleading to call the book 'An Introduction to the Behaviour . . . ', for it deals almost entirely with the correlation between sensation and the anatomy of sense organs. Not only is this a restricted aspect of the whole field, but the survey within it is so comprehensive and detailed that it would require considerable effort for a novice to understand it. As a directory to this field, Dr Carthy has provided a very useful book, a valuable digest and a fruitful source of references: but it is a work for the library rather than the student.

D. M. VOWLES

International Directory of Psychologists Exclusive of the U.S.A. Ed. by E. H. JACOBSON. Assen, The Netherlands: Royal Van Gorcum and Comp. Ltd.; Washington: National Academy of Sciences. 1958. Pp. xxii+527. Hfl. 9.50. U.S. $2.50.

This attractively produced and well-printed volume supplements the familiar *Directory of Psychologists in the United States* by listing some seven thousand psychologists in the various countries of Africa, Asia, Australia and New Zealand, Canada, Europe, Latin America, and the Middle East. Psychologists are listed by countries and there is a convenient name index. Each entry states name and title; preferred postal address; date and place of birth; highest earned degree; University granting this degree and year of graduation; membership of professional societies; editorial responsibilities; current occupation; primary fields of interest in psychology; sex. It is inevitable that such a directory should show omissions and not be uniformly successful in representing psychologists in all communities. But this does not detract from its considerable general usefulness both in facilitating international communication among psychologists and in providing a document of historical interest to future generations. Psychologists owe a debt of gratitude to the many organizations and individuals who collaborated to make possible the publication of this remarkably comprehensive directory.

I. M. L. HUNTER

VOLUME 50 NOVEMBER 1959 PART 4

EFFECT OF PREJUDICE ON REASONING*

BY ROBERT H. THOULESS

Department of Education, University of Cambridge

I. INTRODUCTION

The test I am here describing was first designed as an appendix to the American edition of my book *Straight and Crooked Thinking* and was published in 1939. The publishers had suggested that such a book would sell better if it contained a self-testing device. There were other preoccupations in 1939 and I carried the matter no further then, but I have recently returned to it. My intention in this article is to discuss methods rather than results.

The essential problem of the test was to devise a method by which one could see to what extent different individuals were inclined to judge the logic of an argument as sound if they agreed with its conclusion and unsound if they did not agree with its conclusion. Common experience would lead one to suppose that this is a common tendency and that it is a principle determining much of the effect of propaganda. Can its presence be scientifically demonstrated and measured?

I found in 1958 that the test I devised in 1938 could no longer be used without radical revision. Much of the material of the test had been drawn from controversies that were topical in 1938 but were no longer argued about in 1958—such as fascism and the Italian invasion of Abyssinia. In my revised test, I have still tried to keep the arguments concerned with things that people do really argue about; I think that is essential to any investigation of real problems about the factors determining opinions. So far as possible, however, I have tried to use controversial issues less transitory than those of my earlier test—such as socialism, life after death, and war.

II. EXPERIMENTAL MATERIAL

The general form of the test was the presentation of a series of arguments ending with a conclusion that was underlined. The arguments were in a more or less disguised syllogistic form. For example: 'Slave labour was inefficient because of the lack of incentive to individual effort. Competitive civilizations have prospered because they have provided incentives to individual effort but under socialism such incentives would disappear. We see, therefore, that *socialism is a form of slavery*.'

This, of course, is a fallacious argument of the type the traditional logicians have called 'the fallacy of the undistributed middle'. What I was trying to find out was whether those who disagreed with the conclusion would be more likely to notice that the argument was fallacious than they would have been if they had agreed with the conclusion.

Since I wanted information on the opinions of subjects as to the conclusions at the end of each test argument, the first instruction to the subjects was to read through each of the underlined conclusions while ignoring the rest of the argument and to indicate their agreement or disagreement with it by ringing the word 'true' if he thought it true, 'false' if he thought it false, and to ring a query sign if he had no definite opinion as to its truth or falsity. When this had been done for all test items, subjects were instructed to make no further alterations to the true–false ringings but to go on with the next part of the test, which was to read through the whole of each argument and to ring the word 'sound' if they thought the argument was sound, and 'unsound' if they considered it not sound.

---

* Read at the Glasgow meeting of the British Association for the Advancement of Science, Section J, Thursday, 28 August 1958. The author will be glad to supply a copy of the tests and instructions to any psychologist who wishes to make use of them.

These instructions were distributed to subjects on a typed sheet which contained also an explanation of the exact sense in which the words 'sound' and 'unsound' were to be understood. It was emphasized that the subject was not asked to judge whether the statements of the premisses of the argument were true or false, but only whether, if they were true, the conclusion would follow rigidly from them.

To make this as clear as possible, two examples were given: one of a sound argument and the other of an unsound argument both leading to conclusions that would be generally recognized as true. The examples were:

(1) All fishes breathe under water, but whales have to come up to the air in order to breathe, so we know that *whales are not fishes.*

(2) All fishes breathe under water, and sharks breathe under water, so *sharks are fishes.*

It was pointed out that the first conclusion follows rigidly from what is stated in the earlier part of the argument, while the second conclusion, though also true, does not follow necessarily from what has gone before, so the first argument should be marked as 'sound' while the second should be marked as 'unsound'.

When each subject finished the test, he handed the completed test form back to the experimenter, who marked as right or wrong the judgements of soundness and unsoundness. The primary interest of the experiment is not, of course, in the absolute number of right and wrong answers, but in how the two possible directions of error were distributed amongst the items of which the conclusion was marked as 'true' and those of which it was marked as 'false'.

III. METHOD OF ASSESSMENT

Those items in which the truth of the conclusion was marked as 'doubtful' have no bearing, one way or the other, on the hypothesis under test so they were ignored. There remain eight classes of response made up of the eight possible combinations of the alternatives: conclusion judged 'true' or 'false'; argument judged 'sound' or 'unsound'; and judgement as to soundness or unsoundness of the argument being right or wrong. The eight classes given by the combinations of these possibilities can be most conveniently indicated by the use of three-letter symbols: T or F for 'true' or 'false', S or U for 'sound' or 'unsound', and R or W for 'right' or 'wrong'.

The eight types of response are then: TSR, TSW, TUR, TUW, FSR, FSW, FUR and FUW. Of these, the classes TSW and FUW comprise errors that are in the direction to be expected if there is a tendency to judge arguments as sound if one agrees with their conclusions and as unsound if one does not; the classes TUW and FSW comprise errors in the opposite direction to this. An index showing for a particular subject the influence of prejudice in determining his judgement as to soundness and unsoundness of arguments must be derived from some relationship between the numbers in these classes.

I published an index in 1939 which was altogether erroneous. I cannot now remember by what twisted process of thought I was ever led to suppose it right; I can only try to do better now.

The index I use now is:

$$100 \times \left[\frac{\mathrm{TSW}+\mathrm{FUW}}{\mathrm{TSW}+\mathrm{FUW}+\mathrm{TUR}+\mathrm{FSR}} - \frac{\mathrm{TUW}+\mathrm{FSW}}{\mathrm{TUW}+\mathrm{FSW}+\mathrm{TSR}+\mathrm{FUR}}\right].$$

Expressed in words, this index is the difference between the number of times the subject made an error in the direction of his prejudice expressed as a percentage of the number of times he had the opportunity of making such an error and the number of times he made an error in the opposite direction also expressed as a percentage of the number of times he had an opportunity of making such an error.

This index can be used: (1) For assessing the influence of prejudice in a single subject; (2) for assessing the influence of prejudice in a group of subjects; (3) for assessing the influence of prejudice in the response by a group of subjects to a single test item. In the third case, the formula is simplified by the fact that the item must be either sound or unsound, so only four classes of response are found instead of eight.

IV. RESULTS

The first test of this type which I constructed consisted of sixteen items which were fairly short and not intended to be very difficult. I shall refer to this as test A. Its first application was to twenty-five members of a week-end gathering of adult students with ages ranging from 23 to 58, for whose co-operation I am very grateful. For this group, the test proved to be an efficient instrument for detecting the tendency which I was investigating. The proportion of erroneous judgements was large (about 40%) and of these a large proportion (73%) were in the direction expected (i.e. of thinking that an argument was sound if the conclusion was accepted and unsound if rejected). The individual indices of prejudice went up to 83. The group index was 26·5 which was highly significant

$$(\chi^2 = 24,\ P = 10^{-6}).$$

Later I prepared another test of the same general structure, but intended it to be harder than test A. I shall call this test B. There were twenty items, each of which was considerably longer than the items of test A, since it was thought that one of the factors in disturbing judgement as to soundness of arguments was the concealment of the structure of the thought by the use of many words. Another new factor introduced into test B was the use of attitude-determining words and phrases, also a factor in everyday propaganda. The items were also more closely linked with political and religious problems about which subjects might be expected to feel strongly, since ordinary experience would lead one to suppose that it is in connexion with such problems that errors of the type I was investigating are found. The idea was not to make any separate assessment of the influence of these factors, for which the number of items would have been altogether inadequate, but merely to maximize the total effect in test B.

Applying both of these tests to two groups of students (with thirty in one group and twenty-nine in the other) from the Education Department at Cambridge, results were obtained which were surprising and in some ways disappointing. First, it seemed that both tests were too easy for university graduates. Whereas the adult student group had

had 40% errors in test A, the university group had only 10% errors. For the somewhat harder test B, they had only 16% errors, and five out of the fifty-nine scores was error free. Although this was a smaller number of errors than I had hoped for, it would have been enough to show the tendency I was after if the errors had been distributed in the expected way. In fact, they were not and, except for one subject who had an index of 89% on test A and 20% on test B, there seemed to be evidence of a tendency for the judgement as to soundness to be affected by opinion as to the truth of the conclusion. The only group index that was not negligibly small was for test B in the group which did test B first. This was 7% and is marginally significant ($P = 0{\cdot}03$) although this significance disappears if we take into account the fact that it is the selected best of four group indices.

Some of the individual scores were negative and I considered the possibility that the over-all low group indices were due to some subjects showing the expected effect of prejudice while some had realized the purpose of the test and were reacting in the opposite direction, tending to distrust the soundness of arguments which led to conclusions they believed in because they were aware of their tendency to be misled by their prejudices and were making an effort to counteract that tendency. Detailed examination of the scores did not support this view; there appeared to be only such negative indices as would occur by the chances of sampling. What appeared to be the most striking negative index was that of a subject who had no errors out of eight opportunities in the expected (i.e. prejudice determined) direction and five errors in eleven opportunities in the opposite direction. This gives an index for this subject of −45% which is large enough to tempt one to regard it as not accidental. When this score is reduced to a four-fold contingency table, and its significance is tested by the exact method, it is found that $P = 0{\cdot}04$. This, however, is not evidence that the result is non-accidental, since it is not an isolated occurrence, but the selected most striking score in that direction amongst nearly 100 test results, and the likelihood of a P of 0·04 occurring at least once in so many trials is very high. This apparently striking negative score is, therefore, not evidence of a tendency for some of the subjects to be reacting against the direction of their prejudices. Other methods of examination of the figures confirmed the opinion that, with the exception of the one subject already mentioned, the results of these two graduate groups showed no influence in either direction of the expected effect of prejudice on judgement as to soundness of argument.

There are two further directions of research suggested by this result. First, the difference between the graduate group and the other group is interesting in itself. It is clear that the group of graduates were measurably less inclined to judge the soundness of an argument by their agreement with its conclusion than were the other group. It would be necessary to apply the tests to many other groups before one could discover what were the factors making this difference. It might be a difference in intelligence level, it might be a difference in education, or it might be merely a difference in degree of test sophistication. My results do not enable me to have any opinion as to which is the determining cause. If it turned out to be an effect of difference in education, it would be a surprising and encouraging side effect of university education since, so far as I know, none of the students had studied logic.

The second further question is whether another type of test would yield a different result with the graduate group. I find it difficult to believe that either a good intelligence or a superior education render one altogether immune to the tendency to accept more easily the soundness of an argument if one agrees with its conclusions, although one may under favourable conditions react successfully against that tendency. Such favourable conditions may be provided by the test situation as it has been arranged so far.

First, there is the mere fact that it is known to be a mental test of some kind, and the state of preparedness with which one approaches a psychological test is different from that in which one approaches a political meeting or in which one reads the leading article of a newspaper. This, I think, is one of the most persistent difficulties in making an experimental study of the types of errors in verbal processes which seem apparent to ordinary non-experimental observation. The experimental situation alerts the subject and his thought processes are free from many of the influences which would determine them in the more ordinary situation of listening to a speech or reading a newspaper.

One way of trying to avoid this difficulty is to give the subject a task in which the aim appears to him to be somewhat different from the real aim of the experimenter. In this case, the test looks like a test of the ability to reason correctly and I had some hope that the subjects would regard it as that. A number of those who took it in the graduate group, however, spontaneously asked me after the conclusion of the second test whether the object was not to discover how their opinions about the conclusion of an argument influenced their judgement about its soundness. So, for some subjects at any rate, the real aim of the test was not effectively concealed, and this may have made them less prone to display the weakness in question.

This suggests that, for intelligent and test-sophisticated subjects, the test should be modified so as to make its object less apparent. One obvious modification that I should make in any future revision of the test is that of printing the conclusions on a separate sheet to be marked by the subjects and collected before the sheets containing the arguments are distributed.

Also I think that it was a mistake to give the student group two tests A and B. It would have been better to run these together as a single test with thirty-six items. There is obvious danger that, if two tests are given, the first test may alert the subject as to what the experimenter is after, so that in doing the second test he is on his guard against giving himself away. In fact, both student groups gave a lower prejudice score in the second test, one group having done test B second and the other group test A second. The differences were not significant so no conclusion can be drawn from them, but obviously this is something that might happen and the best safeguard against it would seem to be to make one test and not two.

Also I think it might be well to try presentation of the arguments in speech and not by writing. Most effective propaganda takes place by speech, and heard material may give a better picture of a primary reaction unadulterated by after-thoughts.

To sum up. This method seems to be a promising one for the purpose intended; to measure how far acceptance of an argument as sound is determined by opinion as to the truth of its conclusion. It gave a highly significant indication of the effect in question with one group, but not with another with a different educational background. The next question would seem to be whether there are directions of modification of the test that would lead to it giving measurable results with a group of university graduate subjects.

(*Manuscript received* 8 *September* 1958)

THE ANCHORING EFFECTS OF VALUE IN A SCALE OF JUDGEMENTS

By H. TAJFEL

University of Oxford Delegacy for Social Administration

The investigation is concerned with assessing the effects of value on judgements of weight of a series of stimuli. It was found that an association of value with either the heavy or the light end of the series leads to changes in the judged relationships between the stimuli of the series; the most pronounced effect is that the stimuli at the opposite ends of the series tend to be judged as more different from each other in the value than in the neutral condition. In a control experiment, in which value was associated indiscriminately with all the stimuli of the series, no such effects were observed. Further control experiments, in which the end stimuli of the series were emphasized by means other than attaching value to them, also yielded negative results. The study is viewed in the context of the general problem of the effects that abstract attributes of stimuli may have on judgements of their physical characteristics. An explanation of the findings is formulated in terms of anchoring, and their relevance to perceptual over-estimation and to some related problems is discussed.

I. INTRODUCTION

Most psychophysical investigations are concerned with the effects of various physical attributes of stimuli on judgement. Controlled changes in the attributes of stimuli which affect judgement in various ways are physically defined and measured; these changes are either introduced at the time of judgement, or at least confined within the limits of the experimental session. The concern of other studies, one step removed from the direct alteration of physical attributes at the time of judgement, is with the assessment of the effects on judgement of similar variables, present in the subject's past experience. Here, also, the variables in the subject's past experience which influence his judgements in the present are physically definable and measurable. However, stimuli whose magnitudes we often judge in 'real life' differ from each other not only in their physical attributes. They are often classified according to some abstract criteria; they may be more or less emphasized at the time of judgement because of their different relevance to the situation in which judgement is made, or because of differences between them in the value and meaning that they carry for the subject. These are 'abstract' attributes of the stimuli, in the sense that they cannot be directly observed in the stimuli, either here-and-now or in the past. It can be shown, however, that these abstract attributes do have some effects on the judgement of the physical attributes of the stimuli; that these effects are relevant in a variety of judgement situations (especially social ones): and that they can be predicted with the help of very few assumptions (for a more detailed discussion see Tajfel, 1959).

Experimental studies of the effects of these abstract attributes on judgement are few and far between. The role played by the relevance of a stimulus to the judgement situation was emphasized in a study by Brown (1953). He found that the effectiveness of an anchor may depend on whether this anchor is considered by the subject to belong to the series which he is requested to judge. As Brown puts it: 'In addition to the physical value of the anchoring stimulus, the manner in which the anchor is perceived by *O* may play a role in its effect upon *O*. Above all, it may be crucial whether *O* perceives the anchor as belonging to the other stimuli being judged' (p. 199). Another abstract attribute which may be considered in terms of its 'relevance' is the value to the subject of the stimulus whose physical magnitude he is judging. Many studies, primarily concerned with perceptual over-estimation due to 'motivational' or 'value' variables, threw up as a by-product data which almost invariably pointed to a specific effect of a situation in which changes in value were, in the series of stimuli, correlated with changes in the physical dimension under judgement; this effect consisted of an increase in the judged differences between the physical magnitudes of the stimuli of the series (see McCurdy, 1956; Tajfel, 1957).

The fact that, in the present experiments, a consistent association was established between value and the extremes of the stimulus series leads to implications which go beyond the problem of perceptual overestimation. The question is: can it be shown that endowing some stimuli with an abstract attribute, such as value, would result in certain predictable shifts in the judgement of the *physical* attributes of these stimuli, and of other stimuli belonging to the same series?

Eriksen & Hake were able to show recently (1955, 1957) that their 'subjective-standard' hypothesis predicted better discriminability for those stimuli in a series which their subjects selected as anchors. They argue that normally, in absolute judgement situations, end stimuli of the series tend to be selected as anchors and that better discrimination at the ends of the series is the result of the use by the subjects of these 'subjective standards'. '...Judgements are never made in a vacuum but are always relative to a standard or reference level that is subjectively present. The *S*, when confronted with a series of stimuli in an absolute judgement task, selects a few stimuli that he then uses as standards for judging the remaining stimuli. When a different stimulus is presented *S* attempts to recall one of these 'standard' stimuli and uses it as a reference in judging the presented stimulus. If it is assumed that *S*'s select the end stimuli to use as their subjective standards, the observed anchor effects are the expected outcome. The frequent attempt to recall the value of the end stimuli should increase the accuracy of recognition of these stimuli when they do occur' (Eriksen & Hake, 1957, pp. 132–3).

If it is true to say that abstract attributes, such as value, have an effect on the judgement of the physical characteristics of the stimuli, then an association of value with the extremes of the series may be considered as a special case of the use of 'subjective standards'. This association should make the end stimuli more salient, more apt to be relied upon as 'subjective standards', and the prediction would be that this increase in distinctiveness would result in still further increase in discriminability for the end stimuli, serving as anchors, and for those which are near to them.

The 'relative' or comparative nature of an absolute judgement task is enhanced in a situation where category judgements are used instead of judgements in terms of some unit with which the subject is familiar before the experiment. In the latter case, the judge has a ready-made frame of reference when he enters the experimental situation, and uses it, even though it has been shown that this frame can be considerably modified by a variety of experimental situations. In the case of category judgements, the subject's frame of reference is much more directly determined by the stimuli used in the experiment, and the judgements of the various stimuli of the series are more directly interdependent. Thus, in a category judgement situation a stimulus is 'heavy' only in relation to another stimulus being 'light'.

Therefore, if the assumption is justified that an association of value with the stimuli at one end of the series enhances the anchoring properties of these stimuli, it may be expected that, when category judgements are used, the effects of this enhancement will manifest themselves not only at the end of the series at which value is introduced, but also throughout the entire series, and mainly at its opposite end. A 'heavy' stimulus, consistently paired with some form of value, will become more distinctively heavy in relation to other stimuli of the series, especially to those which, in comparison with it, are most distinctively 'light'. The prediction can thus be made that the introduction of the variable of value at one end of the series will act in the direction of increasing the 'subjective standard' properties of the stimuli at *both* ends of the series.

This has some further implications: Eriksen & Hake (1957) found that accuracy of discrimination was a function of the distance of the stimuli from those which were selected as anchors. In a situation which calls for category judgements, and in which the end stimuli cannot be recognized by the subject with complete certainty, it may be expected that an increased salience of the ends of the series would not just affect these ends. It would, rather, tend to create a more pronounced subjective division of the series into two, or possibly three, 'zones': in the case of weights, for example, into zones of 'heavy' and 'light', and possibly 'medium'.

In the present experiments, judgements of weight were used. The following specific predictions can now be made about the effects of associating value with either of the ends of the series (i) *Range of the response scale.* (*a*) The stimuli at the opposite ends of the series should be judged as more different (or distant) from each other in the 'value' than in the 'neutral' condition. (*b*) This increase in judged distances between the stimuli, most pronounced for the extremes of the series, should be, for the other stimuli, approximately in direct relation to their distance from the middle of the series. (ii) *Shifts in judgements of individual stimuli.* An increase in the range of the response scale may be due to a variety of possible shifts in the judgements of individual stimuli. For example, the whole series of stimuli might be judged heavier, with the stimuli at the heavy end shifting more than those at the light end; or the whole series might be lighter, with the stimuli at the light end shifting more than those at the heavy end; or one end of the scale might remain unchanged, the opposite one shifting in the direction of 'heavier' or 'lighter'. The preceding argument is a commitment to the prediction that the increase in the range of the response scale will be due to the stimuli at or near the heavy end of the series being judged heavier in the 'value' than in the 'neutral' condition, and those at or near the light end being judged lighter in the 'value' than in the 'neutral' condition. The amount of shift should again be inversely related to the distance of the stimulus from the end of the series which is nearer to it.

II. Method

Subjects. A total of 64 undergraduates in the University of Durham were subjects in the four experiments. None of them were reading for a degree in psychology.

Stimuli. A series of10 weights was used: the lightest weighed 200 g., and there was an increase of 10% in weight from one stimulus to the one next it in the order of heaviness in the series. The weights were manufactured by filling identical tin cans with shot.

Apparatus. The subject and the experimenter were seated at two ends of a table with a screen 2½ ft. wide and 2 ft. high between them. The screen was fixed across the middle of a rectangular wooden base screwed to the table at both ends. On this base was fixed a steel bar 1 in. wide and 2 ft. long; this bar, whose centre passed through a narrow rectangular hole in the screen, was evenly balanced about its centre on a steel rod fulcrum raised from the base by 5 in. At the experimenter's end of the bar was fixed a pan on which weights could be placed; on the subject's end was a knob which he was instructed to press downwards as soon as he heard a bell from a timing device placed on another table. The time interval between successive presentations of the stimuli was fixed at 10 sec. Foam rubber pads were fixed on to the wooden base at both ends at the points of fall of the steel bar. The height of the drop of the bar was 8½ in. In this situation, the subject's judgements of weight were given without interference of any visual cues from the stimuli.

Procedure. Each experimental session consisted of two parts. Each subject underwent two experimental sessions at approximately 24 hr. interval.

When the subject entered the room, he was told that the experiment was concerned with the effects of monotony on the fineness of discrimination of weight, and that his task would be to press the knob at his end of the table every time when he heard the bell ringing. In the first part of the session he was to press the bar all the way down to the rubber foam, and then let go and wait until the bell rang again, by which time his end of the bar would be again in its highest position. *No judgements of weight* were to be given by him in this first part of the session. After a few minutes of rest, the second part of the session would start. During this second stage, his task was going to be exactly as before, except that each time as soon as he pressed the bar all the way down, and not before this was done, he was to give estimates of weight. These estimates had to be reported in numbers from one to seven, one standing for the lightest weight, seven for the heaviest.

At this point, each subject was given a few practice trials, announcing judgements of weight, until he and the experimenter were quite certain that instructions had been understood. Before the beginning of the first part of the session, the subject was reminded that no judgements were to be given; before the beginning of the second part, that judgements in terms of numbers from one to seven in increasing order of heaviness should now be reported. No information about the actual number of weights used in the experiments was given to the subjects.

The first part of each session consisted of twelve consecutive randomized presentations of the whole series of stimuli. In the second part of each session, eleven randomized sequences of the series were consecutively presented. These orders of presentation remained unchanged for all subjects in all sessions in all the experiments.

Experiment I

In this experiment, value was associated with the heavy end of the series. This was done by presenting the subject with a bonus worth 3*d*. towards a book token every time when one of the heaviest two stimuli of the series was presented in the *first* part of the 'value' sessions (i.e. the first 12 presentations of the series in a session.) Thus, the subject received 24 bonuses of a total value of 6*s*. The bonus was dropped by the experimenter on the subject's side of the screen while the latter was pressing the bar downwards with one of the two heaviest stimuli on the pan.

Sixteen subjects took part in this experiment, each undergoing two experimental sessions in two consecutive days. They were randomly divided into two groups of 8. One group (the *V–N* group) had value associated with the heaviest two stimuli in the first part of its first experimental session; the following day, in the first part of the second experimental session, the series was presented 12 times as before, but no bonuses were given. For the second group (*N–V*) this order was reversed. In the second part of all experimental sessions no bonuses were presented, and the subjects were requested to report judgements of weight.

At the session during which bonuses were to be given, the subjects were told that 'today, in order to break the monotony of the task, you will be given from time to time paper bonuses'. The worth of the bonuses was stated, and the subjects told that they could exchange them later for book tokens.

This procedure offered three advantages. First, a miniature 'culture' was created in the first part of the value sessions, in which 'heavy' was associated with 'good' or 'valuable': the effects of this association created in the first part of value sessions on judgements of weight reported in the second part of these sessions, during which no bonuses were given, could thus be assessed. Secondly, the procedure offered the possibility of comparing each subject's response scale after the value presentations of the series with his response scale after the neutral presentations. The third advantage was that, in all sessions in all the experiments, the conditions under which judgements were reported (i.e. the second part of each session) were identical. The only difference lay in the subject's immediate past experience, acquired in the first part of their experimental sessions.

Experiment II

A further 16 subjects took part in this experiment. The procedure was as in Exp. I, the only difference being that the bonuses were given to the subjects in the first part of the value sessions at each presentation of one of the two *lightest* stimuli of the series.

Experiment III

This experiment was conducted because of the possibility that positive results in the first two experiments might have been due not to the association of value with the ends of the series, but to the greater attention paid by the subjects to the stimuli and the task (and therefore, possibly, better discrimination) during the value sessions than during the neutral sessions. The procedure remained the same as in the two preceding experiments, except that in the first part of the value sessions bonuses were associated with all the stimuli of the series. The order of pairing the stimuli with bonuses was randomized. The stimuli 1, 3, 5, 7, 8 and 10 were associated with bonuses twice each, and the stimuli 2, 4, 6 and 9 three times each, so as to equate the total amount of reward for each subject with that gained by each subject in the first two experiments. Sixteen subjects were again divided into two groups (*V–N* and *N–V*) in the manner already described for the preceding experiments.

Experiment IV

As should be clear from the Introduction, value is expected to exercise, under conditions described here, certain effects on the response scale not because of some mysterious, undefinable quality of which it is possessed, but because it lends saliency and distinctiveness to the stimuli with which it is associated. The fact that paper bonuses were passed to the subjects at the presentation of the heaviest two or the lightest two stimuli of the series in Exps. I and II might, by itself, have given to these stimuli a certain distinctiveness, independently of whether the bonuses had any value or not. A further control experiment was thus necessary, under conditions identical to those already described for the first two experiments, except that the paper bonuses passed to the subjects would be devoid of all value. In Exp. IV, 16 subjects were divided into four groups as follows. Group 1: order *V–N*; valueless paper bonuses associated with two heaviest stimuli. Group 2: order *N–V*; as above, order of sessions reversed. Group 3: order *V–N*; valueless paper bonuses associated with two lightest stimuli. Group 4: order *N–V*; as above, order of sessions reversed.

Two slight changes were made in the instructions preceding the *V* sessions. First, the subjects were told that in order to break the monotony of the task, they would from time to time receive small pieces of paper. Secondly, in order to keep the value sessions of this experiment as near as possible to the conditions of the value sessions in the first three experiments, the subjects were told that, at the end of the first part of the session, they would receive a bonus worth 6*s.* for a book token. Table 1 sets out the summary of procedure in the four experiments.

Table 1. *Summary of procedure in the four experiments*

	Group *V–N*		Group *N–V*	
	First session	Second session	First session	Second session
Exp. I				
First part of session	AH	FH	FH	AH
Second part of session	FG	FG	FG	FG
Exp. II				
First part of session	BH	FH	FH	BH
Second part of session	FG	FG	FG	FG
Exp. III				
First part of session	CH	FH	FH	CH
Second part of session	FG	FG	FG	FG
Exp. IV (*a*)				
First part of session	DH	FH	FH	DH
Second part of session	FG	FG	FG	FG
Exp. IV (*b*)				
First part of session	EH	FH	FH	EH
Second part of session	FG	FG	FG	FG

A: bonuses for book tokens given with each presentation of one of the two heaviest stimuli of the series. B: bonuses for book tokens given with each presentation of one of the two lightest stimuli of the series. C: bonuses for book tokens given with some presentations of all the stimuli of the series. D: valueless bonuses given with each presentation of one of the two heaviest stimuli of the series. E: valueless bonuses given with each presentation of one of the two lightest stimuli of the series. F: no bonuses given with any of the stimuli. G: judgements of weight required of the subjects. H: judgements of weight not required of the subjects.

III. Results

The over-all mean scale values of stimuli in both experimental conditions in the four experiments are set out in Table 2.

Table 2. *Over-all mean scale values of stimuli in the four experiments*

		Stimuli										Range
		1	2	3	4	5	6	7	8	9	10	1–10
Exp. I	*V*	2·00	2·49	2·87	3·46	4·01	4·55	5·16	5·61	6·29	6·64	4·64
	N	2·24	2·64	2·97	3·59	3·99	4·54	5·19	5·56	6·14	6·50	4·26
Exp. II	*V*	1·67	2·16	2·57	3·03	3·59	4·12	4·91	5·42	5·99	6·50	4·83
	N	1·92	2·38	2·69	3·07	3·64	4·36	4·81	5·41	5·96	6·43	4·51
Exp. III	*V*	1·97	2·24	2·84	3·39	3·89	4·57	5·23	5·57	6·21	6·52	4·55
	N	2·05	2·38	2·81	3·37	3·86	4·46	5·04	5·60	6·14	6·55	4·50
Exp. IV	*V*	2·20	2·51	3·04	3·46	3·94	4·42	5·06	5·44	6·07	6·45	4·25
	N	2·16	2·62	3·09	3·61	3·99	4·53	5·05	5·54	6·03	6·39	4·23

V: Experimental ('value') sessions. *N*: control ('neutral') sessions.

(a) *Range of the response scale*

(1) *Category assignments to the two end stimuli.* It was stated in the Introduction that the most marked effects of the experimental situation described earlier would be expected at the ends of the stimulus series, and that they would consist of an increase in the judged differences between the stimuli lying at the opposite extremes of the series. This was tested by calculating for each subject in each condition the difference between the mean category assignments to the lightest (stimulus 1) and the heaviest (stimulus 10) stimuli of the series (range 1–10).

An analysis of variance for cross-over designs (Cochran & Cox, 1957) was then applied to the scores thus obtained. This was necessary in order to assess the effects of practice, as a comparison of each subject with himself under the value and neutral conditions was built into the design by having half the subjects judging the series first in the value and then in the neutral condition (*V*–*N* order), and the other half in the reverse order (*N*–*V*). The increase in range for the value condition was found significant at the level of $P < 0·01$ in Exp. I and of $P < 0·05$ in Exp. II. Similar analysis applied to Exps. III and IV did not show significant results.

A supplementary assessment of the differences in range 1–10 between the value and the neutral conditions was obtained by comparing these ranges in all the *first* sessions in the value condition in Exps. I and II with those of all the *first* sessions in the neutral condition in all the four experiments (i.e. 16 first sessions in the value condition and 32 first sessions in the neutral condition). The Mann–Whitney U test for large samples (with correction for ties) was used (Siegel, 1956). The results are shown in Table 3.

Similar tests were run to compare the first sessions in: the value condition in Exp. III with all neutral first sessions; the token value condition in Exp. IV with all neutral first sessions; the value condition in Exp. III with value condition in Exps. I and II; and the token value condition in Exp. IV with value condition in Exps. I and II. None of these reached significance.

(2) *Intermediate ranges.* In order to obtain more detailed information about the stretching of the response scale outwards from the middle, distances between judgements assigned to other stimuli of the series were examined 'concentrically'; i.e. scores were

obtained from the data of the judged differences between the stimuli 2 and 9, 3 and 8, 4 and 7, and 5 and 6. In Exps. I and II there was the expected (and very consistent) tendency for the ranges in the value condition to be larger than those in the neutral condition. The differences between the value and the neutral conditions increase as the ends of the series are approached. No such tendency is displayed in Exps. III and IV. (See Fig. 1.)

Table 3. *Range of the response scale in first sessions for value condition (Exps. I and II) and for neutral condition (all experiments)*

V	5·5	5·2	5·2	4·9	4·9	4·9	4·9	4·9	4·7	4·7	4·5	4·4	4·3	4·2	3·4	2·7
N	5·4	5·4	5·3	4·9	4·8	4·8	4·7	4·7	4·6	4·6	4·6	4·6	4·6	4·6	4·6	4·4
	4·2	4·1	4·1	4·0	4·0	3·9	3·8	3·8	3·8	3·6	3·6	3·5	3·4	3·1	3·1	2·9

$U = 161$. $z = 2{\cdot}081$. $P = 0{\cdot}019$.

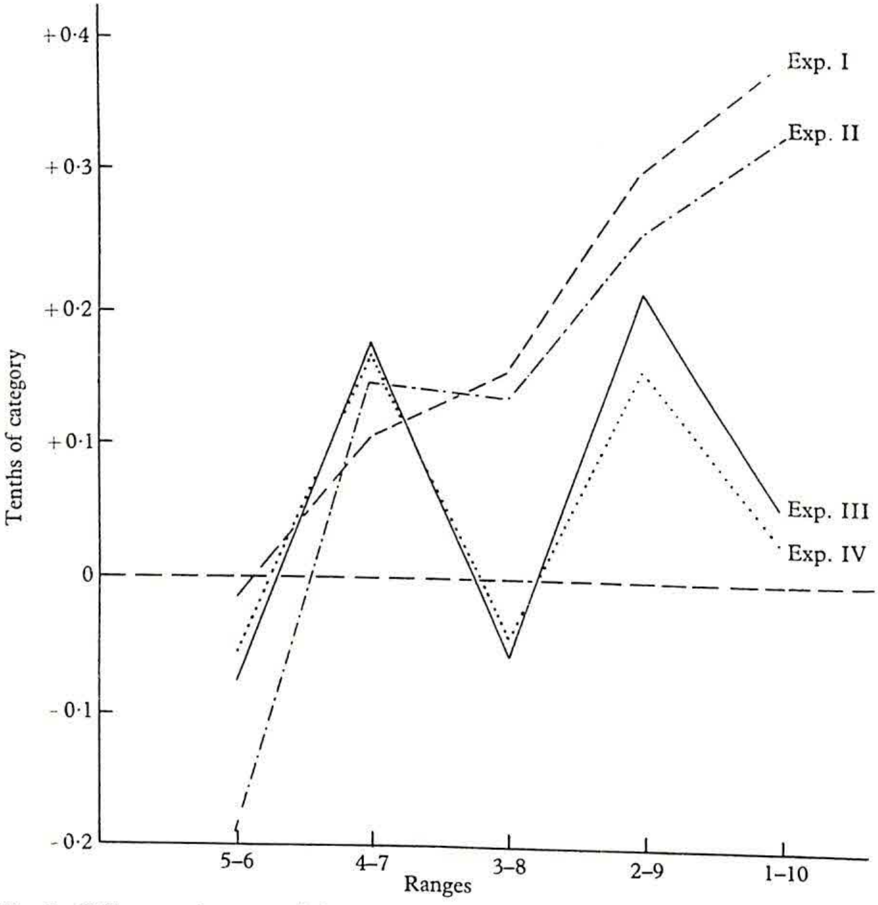


Fig. 1. Differences in range of the response scale between value and neutral conditions.

The graph indicates the *differences* in terms of mean category assignments between the ranges in the value and the neutral conditions. These differences have been plotted as positive when the range in the value condition was larger than in the neutral condition; as negative when the range in the neutral condition was larger.

(*b*) *Shifts of individual stimuli*

As a next step towards a more detailed examination of the nature of the changes in judgements just described, an analysis of variance was carried out for shifts of each stimulus in the series in all the experiments. Two only of these have been found significant:

shifts towards judgements of 'lighter' in the value condition for stimulus 1 in Exps. I and II (P nearly 0·01 in Exp. I; $P < 0\cdot05$ in Exp. II). There is no *a priori* reason why the shift towards judgements of 'lighter' at one end of the series (stimulus 1) should be more marked or consistent than the shift towards judgements of 'heavier' at the other end (stimulus 10) in *both experiments*. The problem raised by this difference between the two ends of the series will be returned to later, in the section devoted to the discussion of results.

(c) *Shifts of the scale*

The results reported earlier for Exps. I and II preclude any marked shifts of the scale as a whole from one experimental condition to another *within* the experiments, as their main feature are shifts in opposite directions of two segments of the series. This was confirmed by an analysis of variance using as data in both experiments the mean scale values for the whole series of stimuli; the results were not significant.

Table 4. *Mean category assignments for the whole series of stimuli: comparison between first sessions in the value condition in Exps. I and II*

Exp. I	4·98	4·75	4·53	4·45	4·33	4·06	3·96	3·82
Exp. II	4·43	4·26	4·14	4·11	3·83	3·71	3·63	2·35

$U = 11\cdot0$. $P = 0\cdot014$.

Table 5. *Mean category assignments for the whole series of stimuli: comparison between first sessions in the value condition in Exp. II and first sessions in the neutral condition in all experiments*

First sessions in value condition (Exp. II):							
4·43	4·26	4·14	4·11	3·83	3·71	3·63	2·35
First sessions in neutral condition (all experiments):							
5·49	4·92	4·89	4·75	4·58	4·56	4·49	4·46
4·41	4·38	4·37	4·36	4·36	4·31	4·28	4·28
4·25	4·24	4·22	4·20	4·03	4·03	4·02	3·92
3·92	3·90	3·87	3·76	3·63	3·55	3·53	2·91

$U = 76\cdot5$. $z = 1\cdot74$. $P = 0\cdot0409$.

Supplementary assessments were made of the differences between the mean scale values for the series of stimuli in the various experimental conditions in different experiments. This was done by running a series of separate Mann–Whitney U tests (Siegel, 1956) for mean scale values in first experimental sessions. The only significant differences were found between: first sessions in the value condition in Exp. I and first sessions in the value condition in Exp. II (in Exp. I the whole series judged 'heavier'; see Table 4); and first sessions in the value condition in Exp. II and first neutral sessions in all the four experiments (in Exp. II the whole series judged 'lighter'; see Table 5). The differences between first sessions in the value condition in Exp. I and first sessions in the neutral condition in all the experiments were in the direction of the series judged as heavier in Exp. I, but they did not reach significance.

IV. DISCUSSION

The experiments confirm the main hypothesis: that an association of value with either end of a series of stimuli results in an increase in the judged differences between the stimuli at or near the opposite ends of the series. This increase seems to be a direct function of the nearness of the stimuli to the ends of the series. It has been argued that these effects of value represent a special case of end-anchoring, i.e. that the association of value with one end of the series increases the salience of the stimuli at both ends of it, and thus enhances their anchoring properties. Some of the results reported earlier are in need of further discussion.

(*a*) *Amount of increase in the range*

It can be seen from Table 2 that the increases in range from the neutral to the value condition in Exps. I and II are in some cases smaller than the absolute differences between the ranges in the various experiments. The significance tests reported in the previous section of the paper have shown that none of these differences in range between the experiments are significant, apart from those (consistent with the hypothesis) between the value condition in Exps. I and II, and the neutral condition in all the experiments (see Table 3). The fact that some non-significant differences in range between the experiments are larger than the predicted differences within Exps. I and II can be understood when it is remembered that the intra-experimental differences are based on judgements reported by the same subjects working under different experimental conditions, and can thus be attributed entirely to the varying of these conditions. The differences in range between the experiments are based on judgements reported by different groups of subjects, and thus represent the initial individual differences in the range in addition to differences due to the various experimental conditions.

(*b*) *Value and discriminability*

One consistent feature of the shifts in judgements of individual stimuli in Exps. I and II was that the shifts towards judgements of 'lighter' of the light end of the series were more marked and consistent than the shifts towards judgements of 'heavier' of its heavy end. The expectation, based on the argument outlined earlier, was that the shifts at both ends of the series would be of approximately equal magnitude, or possibly that a larger shift would occur at the 'value' end of the series. This second alternative seemed more likely.

One explanation of the discrepancy seemed possible: the extent of shifts could reasonably be expected to depend not only on the proximity of the stimuli to the end of the series, but also on the degree of ambiguity of the stimuli. In other words, a subsidiary hypothesis here would be that, at equal distance from the end of the series, the more ambiguous the stimulus, the more markedly will it be shifted. Variability of judgements assigned to a stimulus provides a direct measure of its ambiguity. The simplest way of testing the assumption just stated was to compare the variability of judgements of stimuli 1 and 10 in all those experimental sessions in which the picture was not complicated by irrelevant variables, i.e. the first sessions in the neutral condition in all the four experiments. The S.D.'s of judgements of stimuli 1 and 10 for each subject were calculated, and the Wilcoxon matched-pairs signed-ranks test (Siegel, 1956) was applied to the data. The variability of judgements assigned to stimulus 1 was found greater at a confidence level of $P < 0{\cdot}0002$.

Thus it seems that the greater effect of the experimental conditions on the light end of the series is due to the fact that, with the stimulus range and intervals chosen for the experiments, the heavy end of the series was more clearly and consistently discriminated. This subsidiary finding contains a warning against an over-generalization of the results obtained. Value, and possibly other abstract attributes of stimuli, may be expected to exercise certain effects on the judgements of the physical attributes of these stimuli; but these effects must be expected to be slight when the stimuli are very clearly discriminated. There are, however, three grounds for believing that this limitation of the applicability of the results is not very far-reaching. (i) The range of the response scale was significantly larger for the value condition despite the fact that discrimination between the stimuli of the series was good. Departures from perfect positive correlations between judgements and weights of stimuli in the mean category assignments of individual subjects were very few. (ii) The shifts were more marked for the stimuli at or near the ends of the series despite the fact that the variability of judgements tended to decrease as a direct function of the distance of the stimuli from the middle of the series. (iii) More generally, the usual judgements of magnitude in 'everyday life' can hardly be described as an outcome of slow, cool and mature deliberation undertaken in perceptually optimal and static situations.

(c) *Over-estimation*

The present results are directly relevant to the problem of perceptual over-estimation. 'Over-estimation' due to value or to some related variables has been demonstrated in a number of experiments; but no consistent theory has been offered either to explain it or to predict when it would or would not occur (see Tajfel, 1957). Discussions found in the literature would not allow one to choose between the following alternative predictions of the outcome of the experiments described here. (i) Due to an association of value with the weights, all the stimuli should be judged heavier in the value than in the neutral condition, i.e. there should be an over-all shift towards 'heavier' of the whole series in Exps. I, II and III. (ii) Due to an association of value with certain weights of the series, *these* weights (and possibly the ones near to them) should be judged heavier in the value than in the neutral condition, i.e. in Exp. I the response scale should expand, in Exp. II it should contract. The results do not support either of these predictions.

A variant in the simple statement that an association of value with a stimulus results in an increase of the judged or perceived magnitude of this stimulus exists in one attempt at explaining the phenomenon. This can be roughly summarized as: 'the bigger the better, therefore the better the bigger'. Its corollary would presumably be: 'the smaller the better, therefore the better the smaller'. Murphy's discussions about autism in perception (see Proshansky & Murphy, 1942) would come very near to this formulation. Predictions that would follow from these 'autistic' views would come nearer to some of the results described above: specifically it could be expected that the over-all judgements of weight in Exp. I would shift towards 'heavier'; a shift towards 'lighter' should occur in Exp. II; and no shifts should take place in Exps. III and IV. As was seen earlier, this prediction was not confirmed in an analysis of shifts within each of the experiments; but it found a limited confirmation in comparisons of mean scale values between the experiments (see Tables 4 and 5). On the other hand, this theory has no means of explaining why judgements of some stimuli of the series shifted *away* from the rewarded end. Also, a closer look at the shifts of individual stimuli reveals that in Exp. I the heavier six stimuli of the series shift towards the heavy rewarded end; the lighter four shift in the opposite direction. The reverse happens in Exp. II: the six lighter stimuli shift towards the light rewarded end, the four heavier in the opposite direction. It is quite possible that the differences found between the mean scale values in Exps. I and II reflect no more than this differential pull of the larger segment of the series towards its anchoring rewarded end.

Two further and related points are possibly of some importance. First, since the overall shifts of the series predicted by some form of the autistic hypothesis may be viewed as a by-product of the internal changes within the response scale described earlier, its use as an explanatory and a predictive device conceals the more basic modifications in the judged relationships between the stimuli of the series due to value anchoring. Secondly, the autistic view can hardly be considered as a theory explaining the facts: it is merely restating them in dispositional terms.

My grateful acknowledgements for many helpful discussions are due to: Mr N. T. J. Bailey, Reader in Biometry, and Dr Ian Howarth from the Institute of Experimental Psychology, both in Oxford; Dr R. Audley and Dr A. R. Jonckheere from the Department of Psychology, University College London; and Dr W. Sluckin and Mr Douglas Graham from the Department of Psychology, University of Durham. The investigation was supported by a grant from the Durham Colleges Research Fund.

REFERENCES

BROWN, D. R. (1953). Stimulus-similarity and the anchoring of subjective scales. *Amer. J. Psychol.* **66**, 199–214.

COCHRAN, W. G. & COX, G. M. (1957). *Experimental Designs.* New York: Wiley.

ERIKSEN, C. W. & HAKE, H. W. (1955). Absolute judgements as a function of stimulus range and number of stimulus and response categories. *J. Exp. Psychol.* **49**, 323–32.

ERIKSEN, C. W. & HAKE, H. W. (1957). Anchor effects in absolute judgements. *J. Exp. Psychol.* **53**, 132–8.

McCURDY, H. G. (1956). Coin perception studies and the concept of schemata. *Psychol. Rev.* **63**, 160–8.

PROSHANSKY, H. & MURPHY. G. (1942). The effects of reward and punishment on perception. *J. Psychol.* **13**, 295–305.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York, London: McGraw-Hill.

TAJFEL, H. (1957). Value and the perceptual judgement of magnitude. *Psychol. Rev.* **64**, 192–204.

TAJFEL, H. (1959). Quantitative judgement in social perception. *Brit. J. Psychol.* **50**, 16–29.

(*Manuscript received* 2 *June* 1958)

THE STROBOSCOPIC PATTERNS

II. THE PHENOMENOLOGY OF THE BRIGHT PHASE AND AFTER-IMAGES

By J. R. SMYTHIES*

Psychological Laboratory, University of Cambridge

A detailed study has been made of the bright phase patterns of Brown & Gebhard (1948) produced by looking at a uniform flickering field. The patterns have been classified on the basis of their geometrical properties. The subjects have been classified on a basis of the different types of pattern they report. The types of movement encountered and the effect of varying a number of parameters were investigated. The after-images produced by this means were observed in a similar manner.

I. Introduction

In a previous paper in this *Journal* (Smythies, 1959) an account was given of the stroboscopic patterns (the sensory patterns observed in the visual field on looking at a large uniform surface illuminated by a flickering light) and particular attention was given to the dark phase of these patterns first described by Brown & Gebhard (1948). The present paper will be concerned with a description and analysis of the bright-phase patterns and of the interesting after-images obtained by this means. The bright-phase patterns have been known for considerably longer than have the dark-phase patterns, having been discovered by Purkinje in 1823. He obtained the effect by looking up at the sun with his eyes closed and passing the spread fingers of one hand rapidly to and fro in front of his eyes. He described primary and secondary forms, the latter developing out of the former. Primary forms included a checker-board covering the whole field, large hexagons, shifting zig-zag lines, a central rose-leaf pattern: secondary forms included an eight-rayed star and a rectangular spiral. These were very changeable. Helmholtz (1924) also investigated the phenomenon and reported a central rosette figure surrounded by irregular spots increasing in size to the periphery. After Helmholtz a number of investigators studied the Prévost–Fechner–Benham colours obtained by such means as Benham's top; some of these made brief reference to hexagonal figures, grids, radial lines and mosaics produced by such stimulation (cf. Fechner, 1838; Brewster, 1861; Aubert, 1865; Exner, 1870; Pauli & Wenzl, 1924). Edridge-Green (1925), in his *Physiology of Vision*, described various phenomena induced by intermittent light, such as networks, star figures and hexagonal patterns, all of which he attributed to the entoptic visualization of various retinal structures such as blood vessels, macular pigment or the pigment cells of the retina. This constitutes the first hypothesis of their origin. In recent years Grey Walter (1950) has given a brief description of the phenomenon as encountered in the course of photic stimulation in electroencephalogram work. He noted their geometrical form (particularly checker-boards and catherine wheels) and their incessant movement. He further constructed the second hypothesis as to their origin: that they arise from the interference with a scanning system in the nervous system by the intermittent stimulation and he gives the existence of the patterns as support for his hypothesis that there are such scanning mechanisms in the visual nervous mechanisms. Gebhard (1943) described the patterns in a paper devoted to the chromatic phenomena produced by intermittent stimulation of the retina as follows. 'There were also a variety of mobile geometrical patterns too numerous to mention. These filled the entire central portion of the field and were present most of the time.' In a later paper, Brown & Gebhard (1948) made the first classification of the phenomenon into its 'bright' and 'dark' phases. They noted 'a most remarkable and beautiful display of colour and form perceptions'—the patterns were thought to be roughly similar to Purkinje's descriptions but important differences were noted. The bright phase was dominated in their two subjects by chromatic geometrical patterns—either windmill vanes, a checker-board, hexagons or rainbow bands. The dark phase was described in part I of the present paper. They raised the question: 'Why is the number of different patterns apparently limited when one might have expected limitless variety?' And they concluded: 'These visual patterns are manifestations of intrinsic bias or anisotropy of the visual mechanism, more probably physiological than physical.' Costa (1953) described three types of pattern obtained by sitting with shut eyes in front of a flashing lamp. These types were (*a*) autoscopic, (*b*) geometrical, and (*c*) hallucinatory. (*a*) These patterns were fragmentary and were produced only by very intense lights. They consisted of 'ramifications of retinal veins, hexagonal cells of the choroid, a lozenge probably corresponding to the point of exit of the optic nerve, stripes, circles, etc.' (*b*) The geometrical patterns were predominant and consisted of patterns with radial symmetry or, more rarely, of regular stripes or waves. In some cases, there were strong movement effects with vortices and wavelike motions of the whole field, kaleidoscopic effects, ornamental designs, roman roses and gothic designs. (*c*) A few subjects saw objects such as horses, sails, eyes, spiders, etc. Costa also reported individual differences between subjects. In some, the geometrical patterns were achromatic; in most, they were chromatic; in others, the movement effects were predominant. Mundy-Castle (1953) reported the patterns and noted that 'the most common are visual in the form of moving concentric rings with a faint cross or regular spoked figure radiating from the centre, together with myriads of small shadows or criss-crossed elements forming a shifting network of colours changing both spontaneously and with different stimulus frequencies. At low frequencies (below 8 f./s.) there is less movement and vivid colours seem to fill the head, whilst above 30 f./s. a homogeneous blue-white field is seen. Every subject we have tested (approximately 1000) capable of doing so has reported perception of visual movement, pattern and colour during stimulation within the 4–25 f./s. range'. He put forward the third hypothesis to account for their origin: 'the interaction between spontaneous and evoked electrical rhythms and the augmentation of activity harmonically related to the stimulus frequency'. Blum (1956), in the most extensive study of the bright phase made to date, conducted experiments to see if there was any difference in the type of pattern between normal, brain-damaged and schizophrenic people. During the course of this, he reported the following types of pattern: checks, mesh, cubes, boxes, squares, diamonds, triangles; rays, snowflake-like; ovals, disks, holes, circles, spirals, half-circles, arcs, crescents, tubes, cones, waves; lines, dots, spots, mottle; also, in some subjects, meaningful images.

---

* This work was completed at the Galesburg State Research Hospital, Galesburg, Illinois. These experiments were conducted during the tenure of a Nuffield Fellowship in Medicine.

The present investigation was undertaken with two main objectives. First, to discover the natural history of the patterns in as great a detail as possible. Secondly, to use the form of the patterns under various conditions as a basis for deduction about certain principles of action of the visual mechanisms. The argument here may be put as follows. We can treat the situation essentially in terms of a 'black box'. The input is a temporally intermittent and spatially uniform light stimulus of the retina. The output is a report by the organism of the perception of geometrical patterns (temporally continuous, spatially diversified). That is to say, a faulty *classification* is made of this particular stimulus type. It is therefore possible that, if we can find out any regularities and principles underlying this faulty classification, this will give us direct evidence from which we can deduce, if we know the sorts of classification mechanism that are possible (cf. Uttley (1954) and Beurle (1956)), features about the visual mechanisms themselves. This faulty classification might arise out of faulty, ambiguous or misleading coding by the peripheral mechanisms, or from defects in transmission of the coded messages, or from the inability of the cortical classificatory mechanism to handle these particular configurations of coded messages correctly because of its particular specifications. This argument will be elaborated later, but an example here will clarify the possible gains from such an investigation. Suppose that we have a television mechanism employing scanning, and the problem is to find out the form of the raster employed without inspecting the screen closely. This may be done by asking for the studio to be illuminated by an intermittent light at various frequencies. As soon as the flash frequency falls below the frame frequency, 'illusory' geometrical patterns will appear on the screen in the manner described by Grey Walter (1950). Now the form of these patterns will be determined by the form of the raster used, i.e. a linear raster will give either a series of straight parallel lines vertically orientated in the case of a vertical raster and horizontally orientated in the case of a horizontal raster, or else to patterns built out of short parallel straight lines. Similarly, a polar scan will give rise to circles or sections of circles or spirals or patterns built out of these, and a radial scan will give rise to radial arrangements. Now if Grey Walter is correct in supposing that scanning processes are important in the visual mechanisms and determine the form of the stroboscopic patterns, clearly these forms will give us information as to the form of the raster—or the physiological equivalent of a raster—actually employed by the visual mechanisms. Even if we reject the scanning hypothesis, the patterns may, nevertheless, give us information as to details of operation of the mechanisms responsible for their production. So the details of the patterns become of general interest.

II. MATERIAL AND METHODS

These were the same as described in part I of this paper. The following shorthand will be employed in this paper. The runs with both eyes open after both eyes have been shut will be called *SO*; the runs with one eye open after both have been closed, *SR*; the runs with one eye open after both eyes have been open, *OR*; and the runs with both eyes open after one has been shut, *RO*. In the former three types of run all patterns reported for the whole 60 sec. of the run were counted. In *RO* runs, only the initial patterns were counted. The term *initial pattern* refers to any pattern reported during the first 10 sec. of the run, and the term *developed pattern* refers to any pattern reported thereafter. *SR*, *OR* and *RO* runs refer equally to either eye.

III. RESULTS

(1) *General features of the patterns*

(i) The general features of the patterns experienced by one subject will remain constant for that subject for years. (ii) There is a good deal of individual variation in the patterns and sequence of patterns experienced by different subjects, although this variation is always within definite limits. Subjects may be classified into different types on a basis of these differences (see § (2) below). (iii) Individual differences may be distinguished and classified on the following bases. (*a*) the main types of pattern present and the number of different types reported; (*b*) the sequence in which these patterns occur; (*c*) the complexity of the patterns; (*d*) the constancy of the patterns, i.e. if one pattern tends to stay for a long time or if there is a continual flux and change of pattern; (*e*) the amount of patterning present, i.e. if there is always (at adequate levels of illumination and frequency) a definite pattern present, or if there is just a confused or uniform field most of the time out of which fleeting patterns may or may not appear from time to time; (*f*) whether the patterns fade or disappear on continued stimulation or not; (*g*) the range of response to different parameters of stimulation (illumination, frequency, colour); and (*h*) the differences between binocular and uniocular stimulation.

(2) *Classification of individual subjects*

Method of classification. The patterns observed were either constructed of straight lines or of curved lines and only very rarely of both as connected elements of the pattern. Furthermore, some people observe a few dominant constant patterns, others many changeable patterns. These two facts were used as a basis of the two main types of classification employed. The reports of each subject of each type of pattern were counted and tabulated for both the initial patterns (I.P.: obtained in the first 10 sec. of stimulation) and the developed patterns (D.P.: obtained after the first 10 sec.). Points of detail here were as follows: single lines (other than a clear-cut circle, spiral, parabola or other nameable curve) did not count, nor did mere diffuse spots, blobs, mottles, streaks, etc. A central blob or disk was not counted if unpatterned or unmoving, as this might merely represent the entoptic visualization of the macula pigment. One type of pattern was considered dominant if there were more than twice as many examples of it as the alternative. The frequency of patterns was further tabulated into binocular and uniocular stimulation and into white and monochromatic light.

Classification. Subjects may be classified into seven types on a basis of the amounts and relationships of straight-line and curved-line patterns as follows. Class I*a*. Dominant initial straight-line patterns going to dominant curved-line patterns (including central rotating catherine wheels) (7 subjects). I*b*. With binocular stimulation as for I*a*: uniocular stimulation shows very few curved-line patterns (due to the dominance of parallel vertical lines) (3 subjects). I*c*. Equal straight- and curved-line patterns going to dominant curved-line patterns (3 subjects). I*d*. Dominant straight-line patterns going to equal straight- and curved-line patterns (6 subjects). Class II. One particular straight-line pattern alternating strongly with a particular curved-line pattern (3 subjects). Class III. Very few straight lines (1 subject). Class IV. Very few curved lines (2 subjects). Class V. Very variable patterns in no clear sequence (2 subjects). Class VI. Very little pattern (3 subjects). Class VII. Dominant complex maze/mosaic/snowflake patterns (2 subjects). Two subjects remained unclassified (for details see below).

The second means of classification yields three grades. Grade A. Those subjects reporting a large number of different patterns (13). Grade B. Those subjects reporting a few dominant patterns (16). Grade C. Those subjects intermediate between these two (3) (Subjects 4, 22, 29). Of the two subjects unclassified above one was in grade B and one in grade B with the Aldis lamp and in grade A with the strobe.

The details of the patterns found in grades A and C are given in Table 1, and the details of the proportion and distribution of the straight-line and curved-line patterns in Table 2. Brief accounts of those subjects showing a few dominant patterns will be given individually. These tables show that, for grades A and C, radiating lines, grids, squares, parallel lines or circles are *frequent* initial patterns: *moderately frequent* is flower pattern; *moderately infrequent* are cross, diamond, mosaic, ellipse; the rest are *infrequent*. (Scores: 0–10 infrequent: 11–25 moderately infrequent: 26–29 moderately frequent: 40 frequent). Developed patterns: *frequent*, radiating lines, circle, vortex; *moderately frequent*, cross, grid, squares, parallel lines, spiral; *moderately infrequent*, cross, diamond, mosaic, flower pattern; the rest are infrequent.

Grade B. Class I*a*. All patterns observed are given, and in order of frequency (straight-line patterns ordered separately to curved-line patterns).

Subject 3. I.P. A star (8 limbs) or a cross (4 limbs) made up of piece of herring-bone (Pl. 1, fig. 1,.
D.P. Two-banded concentric rotating circles (Pl. 1, fig. 2) (binocular); shimmering centrifugal motion (uniocular). Total species of pattern reported: 3.

Subject 5. I.P. *SO*, a grid or checker-board; *SR*, nil; *OR*, radial lines; *RO*, a grid or mosaic.
D.P. *SO*, a catherine wheel or a central flower pattern rotating; *SR*, *OR*, little bright phase, a dominant dark phase. Total 6.

Subject 9. I.P. A star (Pl. 1, fig. 3) or simple snowflake. *RO*, a check or checker-board.
D.P. A central vortex with the initial pattern lingering on in the periphery of the field as a mottled effect. Total 5.

Subject 10. I.P. *SO*, *RO*, mosaic; *SR*, *OR*, vertical lines.
D.P. A central vortex or spot (binocular); only a central spot (uniocular). Total 3.

Subject 20. I.P. A herring-bone or honeycomb, occasional star, mosaic, circle or spiral.
D.P. A catherine wheel with dim herring-bones in the periphery. The latter may invade the whole field at times in streaming bands; occasional star, cross, mosaic or spiral. Total 8.

Subject 23. I.P. *SO*, *SR*, an initial uniform field for a few seconds in which a vague mosaic develops mainly in the periphery; *OR*, *RO*, the mosaic develops at once.
D.P. A central spiral or catherine wheel: dim herring-bones or mosaic in the periphery. Total 4.

Class I*b*

Subject 11. I.P. *SO*, a central sun with radiations, an occasional diffraction pattern or maze; *SR*, *RO*, a uniform field or occasional radiations; *OR*, a diffraction pattern (Pl. 1, fig. 4), an occasional maze or honeycomb.
D.P. (binocular). A 'chain and vortex pattern' described more fully in § (3): (uniocular) diffraction pattern or radiating sun alternating with a strong dark phase, never the 'chain and vortex pattern'. Total 5.

Class II

Subject 2. I.P. A number of figures variously described as 'iron crosses', 'snow flakes', 'ice crystals' and 'teazle heads' (Text-fig. 13) being variations on a common theme: very occasionally cross, diamond, herring-bone. *OR*, a quite different effect being one of a class of closely allied patterns called 'ripple effects', 'graining effect', 'porphyry', also a molecule effect (Text-fig. 13).
D.P. (binocular). A herring-bone pattern alternating strongly with a central rotating ball; (uniocular) a uniform field or a general ripple with an occasional trace of herring-bone and very dominant dark phase patterns. Total 5.

Subject 8. I.P. A uniform field, once only a cross. *OR*, mandala, cross, circles, spiral (Pl. 1, fig. 5), pseudosphere and vortex.
D.P. A strong and regular alternation between a mandala (Text-fig. 14) and a spiral (Pl. 1, fig. 5); much less frequently a great monad, a snowflake, 'lines of force around a bar magnet' pattern, a pseudosphere, ellipse, doughnut and 'three-legs of the Isle of Man'. Total 12.

Table 1. *Analysis of frequency of type of pattern in grades A and C*

(Columns B, D, number of subjects in which pattern was not seen; B, I.P.; D, D.P. Columns A, C, scored ranking; A, I.P.; C, D.P. This score was obtained by tabulating for each pattern the number of times it was the most frequent pattern reported by a subject, and the second, third and fourth most frequent. A first was scored 8, a second 4, a third 2 and a fourth 1. Notice the preponderance for I.P. of, in order; squares, radiating lines, grid, parallel lines, and circle and flower patterns; for D.P.; vortex, circle, radiating lines, spiral, parallel lines, and squares (the first six is given in each case).)

	I.P.		D.P.	
Pattern	A	B	C	D
1	51	3	42	7
2	14	10	21	7
3	4	13	1	14
4	0	15	0	16
5	5	12	1	14
6	49	4	29	9
7	52	4	32	7
8	11	8	22	7
9	0	16	0	16
10	0	12	8	13
11	44	5	34	10
12	0	15	1	12
13	6	12	0	15
14	2	12	0	16
15	5	10	6	12
16	18	7	17	9
17	0	16	0	16
18	44	8	76	4
19	14	13	6	13
20	0	15	8	15
21	8	14	37	7
22	8	14	8	15
23	9	14	6	13
24	0	16	0	16
25	0	16	0	16
26	2	15	80	2
27	32	13	12	13
28	9	14	0	16

Code for numbering patterns. For tables 1 *and* 3

Straight-line patterns

1. Radiating lines
2. Cross
3. Cob-web
4. Mandala *
5. Snowflake
6. Grid
7. Square or rectangle
8. Diamond
9. Diffraction pattern *
10. Checker-board
11. Parallel lines
12. Triangle
13. Herring-bone
14. Zig-zag lines
15. Honeycomb
16. Mosaic
17. Crystals

Curved-line patterns

18. Circle
19. Ellipse
20. Parabola
21. Spiral
22. Hyperbola
23. Sine waves
24. Magnetic field
25. Wedges (*)
26. Vortex
27. Flower pattern
28. Fingerprint pattern

Miscellaneous

29. Design
30. Image
31. Blob, dots, etc.
32. No pattern
33. Miscellaneous straight-line pattern
34. Miscellaneous curved-line pattern
35. Veins

* More than 90% of the total count contributed by 1 subject.

Subject 27. I.P. Concentric circles, central rotation or a cross; rarely vertical or horizontal lines, a honeycomb of wavy lines.

D.P. A cross alternating regularly with a catherine wheel or less often with concentric circles; sometimes one of these alone; rarely radial lines, vertical lines, small squares. Total 8.

Class III

Subject 14. I.P. *SO*, mostly no pattern; occasional concentric circles, circles enclosed in diamonds, honeycomb and the figure shown in Text-fig. 15. *SR*, mostly no pattern; occasional concentric circles, a complex scalloped pattern, a wedge pattern and vertical lines. *OR*, the dominant pattern was a remarkable pattern composed of wedges described further in § (3) (*m*) below. *RO*, the interesting feature here was a spot which appeared in front of the newly opened eye and which took up to 30 sec. to centre; otherwise mostly a uniform field or occasional concentric circles, or the Text-fig. 15 pattern again.

D.P. The dominant pattern was concentric circles sometimes complicating to a flower pattern or complex scalloped pattern (Pl. 1, fig. 6); also magnetic lines of force and the wedge pattern (*SR* and particularly *OR*). Total 7.

Table 2. *Analysis of straight-line patterns* (*SLP*) *and curved-line patterns* (*CLP*) *in each class; Aldis, Strobe and colour*

(In the ratios decimals are only given below 10.)

			I.P.			D.P.		
Class	Subjects		*SLP*	*CLP*	Runs	*SLP*	*CLP*	Runs
I A	7	Ratio	13	1	1237	1	2·4	915
		Range	22–205	0–12	—	0–28	19–110	—
I B	3	Ratio	15	1	296	2·6	1	232
		Range	40–128	2–10	—	44–66	14–27	—
I C	3	Ratio	1·2	1	342	1	4·1	248
		Range	12–41	11–36	—	39	21–68	—
I D	6	Ratio	13	1	795	1·3	1	600
		Range	25–77	2–5	—	11–40	12–26	—
ΣI	19	Ratio	8·1	1	—	1	1·3	—
ΣII	3	Ratio	5·6	1	508	1·2	1	379
		Range	20–82	7–20	—	27–74	13–88	—
ΣIII	1	Ratio	1	5·6	151	1	39	113
		Range	11	62	—	2	79	—
ΣIV	2	Ratio	248	1	347	10	1	276
		Range	75–173	0–1	—	70–77	0–14	—
ΣV	2	Ratio	1·3	1	172	1·9	1	130
		Range	13–53	9–40	—	15–45	9–34	—
ΣVI	3	Ratio	5	1	325	1	1·5	238
		Range	4–26	1–6	—	0–10	4–13	—
ΣVII	2	Ratio	73	1	190	7	1	158
		Range	26–47	0–1	—	26–32	0–8	—
ΣI–VII	32	Ratio	6	1	4363	1	1·05	3289

Class IV

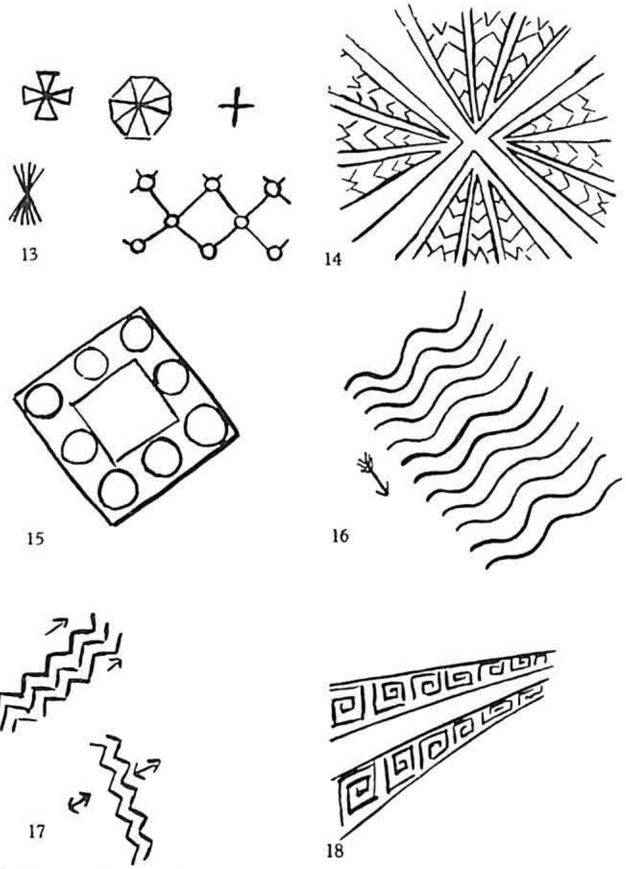
Subject 1. The patterns reported by this subject were remarkable for their interesting and regular form and their regularity of reproduction. *SO* I.P., a pattern of thin blue scroll line in fourfold radial symmetry (Pl. 2, fig. 7); sometimes the thin blue lines formed a star or series of stars instead. D.P., A central diamond with a collection of red and green dots inside and a checker-board in various colours around. *SR*, a similar series except the checker-board came in earlier and the patterns were less well marked. *OR*, here the initial pattern was a checker-board which faded usually to a uniform field sometimes with a diamond and dots. *RO* I.P., the same symmetrical pattern of thin blue lines. Total 4.

Subject 13. I.P. *SO*, *SR*, a diamond spiral (Pl. 2, fig. 8) or central diamond; less often a cobweb, checker-board, four-leaf clover, or a pattern of squares. *OR*, a mosaic or grid; less often a cobweb or a reticulate pattern; *RO*, a checker-board, diamond spiral, diamond, grid, or concentric squares.

D.P. The diamond spiral alternating at times with a group of parallel sine waves arranged diagonally and moving downwards across the field (Text-fig. 16). Total 10.

Subject 24. I.P. A honeycomb (occasional mosaic or cobweb) or circles (occasional vortex or spiral).

D.P. Honeycomb or circles. Total 6.



Text-figs. 13–18. Line drawings copied by the author from line drawings by the subjects made immediately after observation.

Text-fig. 13. 'Iron crosses', 'snowflake' and 'teazle-head' patterns. I.P., subject 2. Also one molecule effect (*OR*). Text-fig. 14. Mandala figure. D.P., subject 8. Text-fig. 15. Complex figure. D.P., subject 14. Text-fig. 16. Sine waves. D.P., subject 13. Text-fig. 17. Types of movement of zig-zag lines. Text-fig. 18. 'Faults' with grecian scrolls. Subject 18.

Class VII

Subject 25. The patterns were dominated by a snowflake formation made up almost entirely of straight lines (thus this subject also belongs to class 4).

I.P. Snowflake with occasional variants of this (butterfly-wing, crystals, Christmas tree); occasional 'falling snow', cross, mosaic, star.

D.P. Snowflake with the same occasional variants; occasional grid, cross, circles, rotation. Total 8.

Subject 35. I.P. A complex line figure of overlapping hexagons in fourfold radial symmetry not covering the centre of the field.

D.P. A complex maze/mosaic pattern (Pl. 2, fig. 9). No other patterns. Total 2.

Unclassified subjects

Subject 6. I.P. and D.P. *Aldis*. The only pattern seen was a cross on a ground bearing an hexagonal pattern (i.e. class IV). Total 2.

Strobe. I.P. Cross, parallel lines, grid, honeycomb, magnetic field pattern; flower pattern, vortex.

D.P. Star, cross, grid, vertical lines, checker-board, diamond spiral; but mostly curved-line patterns—vortex, spiral, ellipses, circles, nebula.

These patterns were sometimes made up of dots but more often by lines (i.e. class I). Total 15.

Subject 7. This subject showed very few patterns with binocular stimulation, mostly just small elliptical disks covering a small area at the centre of the field. With uniocular stimulation a tartan or grid pattern was dominant. Otherwise only very occasionally honeycomb or squares. Total 4.

(3) *Particular features of the patterns*

Classification of the patterns. Class 1. Unformed elements such as odd blobs, splodges, dots, streaks, mottles. Class 2. Single lines, usually diagonals, may cross the field: there may be rainbow lines. Class 3. Patterns based on parallel straight lines. Class 4. Patterns based or a radial arrangement of straight lines. Class 5. Patterns based on straight lines not simply parallel such as a honeycomb, herring-bone, maze or zig-zags. Class 6. Patterns based on curved lines. Class 7. Designs and formed images. The details are given in Table 3.

Classification of movements. (i) Rotation of whole or part of the pattern either clockwise or anticlockwise. Sometimes the centre may rotate one way and the periphery the

Table 3. *Total number of reports of each pattern per* 1000 *runs: white and monochromatic light*

(Total runs: I.P. 4363, D.P. 3289. See Table 1 for code for numbering patterns.)

Pattern	I.P.	D.P.	Total
1	78	34	112
2	58	50	108
3	7	3	10
4	4*	17*	21*
5	15	6	21
Σ	162	110	272
6	37	29	66
7	34	22	56
8	11	34	45
9	7*	17*	24*
10	20	15	35
11	56	32	88
12	1	1	2
Σ	166	150	316
13	4	9	13
14	3	2	5
15	16	11	27
16	50	36	86
17	2	1	3
Σ	75	59	134
18	20	63	83
19	1	2	3
20	0·4	0·3	0·7
21	1	33	34
22	1	1	2
23	1	3	4
24	1	1	2
25	7*	6*	13*
26	12	156	168
27	10	6	16
28	1	0·3	1·3
Σ	55·4	271·6	327
29	1	1	2
30	1	3	4
31	247	276	523
32	286	240	526
33	5	1	6
34	6	8	14
35	1	0·5	1·5
ΣS.L.P.	408	320	728
ΣC.L.P.	61·4	279·6	341

* More than 90% of reports from one subject only.

other. The rotation is usually about a stationary centre but epicyclic rotations also occur; (ii) centrifugal and centripetal motion; (iii) linear translations—particularly of parallel straight or wavy lines or of the whole of a mosaic pattern moving either regularly across the screen or as an irregular jiggle; (iv) rocking of a star figure around its centre or a waving of its arms up and down; (v) a streaming motion along lines, particularly diagonals; (vi) the oscillation of a point; (vii) the relative line-to-line motion of the elements of a zig-zag pattern, either a concertina motion or a creep (Text-fig. 17); (viii) topological stretchings as though the whole pattern were on a rubber sheet being stretched in various directions; (ix) 3-dimensional movements of patterns rotating into the plane away from the observer, or the whole pattern may approach or retreat from the observer.

Interrelation of pattern and movement. Each pattern shows a characteristic range of different movements as follows: (i) *Vortices:* simply rotate either way, or rotate and expand with new vortices appearing at the centre to expand in turn; they may jig up and down. (ii) *Concentric circles:* centrifugal or centripetal progression, rotation and distortion; the centre may itself rotate in an orbit. (iii) *Ellipses:* rotation, the centre may itself rotate; once a curious rocking about one point on the circumference. (iv) *Spiral:* rotation, the centre itself may rotate, distortion. (v) *Cross and star:* rocking, rotation, internal movement along the constituent lines which themselves remain stationary. (vi) *Concentric diamonds:* rotate only, the central ones may rotate in the opposite direction to the peripheral ones. (vii) *Concentric squares:* rotation, centrifugal motion of squares. (viii) *Diamond spiral:* internal motion of dots along its constituent arms, distorts. (ix) *Mandala:* waves its arms, distorts. (x) *Parallel vertical straight lines:* internal movement along the lines, slight displacements and tiltings; they never rotate and only very rarely move at right angles to their own length. These lines in any number (i.e. more than 4–5) are only obtained with uniocular stimulation. (xi) *Parallel straight diagonal or horizontal lines:* these commonly move at right angles to their own length. They are commonly seen with binocular stimulation. (xii) *Parallel vertical, diagonal or horizontal sine waves:* commonly move at right angles to their own length. (xiii) *Grid and checker-board:* these are usually stationary but may move as a whole across the screen. Never rotate. (xiv) *Honeycomb:* rotation, bodily motion across the screen, or two superimposed honeycombs may counterrotate. (xv) *Herring-bone:* this may undergo complex motions—concertina-like, rotation or linear translations of the whole or part, distortion, 3-dimensional movements. (xvi) *A regular mosaic of small squares, triangles, diamonds, etc.:* again this may undergo complex movements: centrifugal and centripetal (often combined with rotation), rotation, linear translations of part or whole, a kaleidoscopic jumble. (xvii) *An irregular labyrinth type of mosaic:* shows rotation, linear and complex translations, rocking, concertina-like movement, relative line-to-line creep.

The complex motions of these last three types of topologically disconnected patterns stand in interesting contrast to the relative paucity of movement of the topologically connected grids and checker-boards. This list is not necessarily exhaustive and shows only what has been observed in 35 subjects.

The modes of change of pattern. One pattern may change into another in any of the following ways. (i) By continuous transformation. Examples which occurred were a family of parabolas which curled round to form concentric circles, which in turn developed into a spiral. Another instance was horizontal lines which developed into a grid and then into a brick-wall effect. These transformations were not very common. (ii) By a sudden change of one to the other without continuity. (iii) The first pattern may break up into a moving mass of broken pattern out of which the new pattern slowly crystallizes. (iv) A patternless field may intervene between the two patterns. A vortex may change its direction of rotation in one of four ways. It may slow down, stop and then start up again the other way; the change may be quite sudden; or the vortex may turn over into the third dimension and thus come back rotating the other way. Once the vortex developed into a figure of eight. The original half then faded leaving the new half as a vortex going round the other way.

Fatigue effects. Two effects of prolonged viewing were noted. In some subjects there was quite a marked fading off of the patterns even after only 30–45 sec. of viewing. In other cases there was no such fading off even after many minutes. Furthermore, if specific effects such as the production of parallel vertical lines with *OR* are sought repeatedly, the pattern will usually be obtained clearly the first few times and then with increasing difficulty. In the experimenter, prolonged viewing induced strong concertina movements.

Periphery versus centre of field. There is very often a difference between the periphery and the centre of the field. The commonest is to have a vortex or some curved-line or rotating central pattern in the centre with a mosaic or herring-bone in the periphery (Pl. 1, fig. 5). Also, there may be a pattern only in the centre or only in the periphery. The size of elements of a pattern may also increase to the periphery (e.g. the diffraction pattern of subject 11 shown in Pl. 1, fig. 4); on the other hand, it may not as in checker-boards or grids which remain uniform.

Uniocular versus binocular stimulation. Most subjects reported that the patterns with uniocular stimulation were generically the same as those seen with binocular stimulation but that they were on the whole much less clear-cut and definite to the extent that definite elements of the pattern might be lost. Two subjects however reported that they were more clear-cut with uniocular stimulation. Subject 7 had very few patterns with binocular stimulation but quite a good 'tweed' or grid pattern with uniocular. Subject 31 showed the usual geometrical patterns of grade C in no clear sequence (class 5) but was definite that they were clearer with uniocular stimulation. The dark-phase patterns and the numerous parallel vertical straight lines occurred only with uniocular stimulation.

The effects of dark and light adaption. The effect of 30 sec. dark adaption is given as the I.P.'s of *SO* and *SR*, and these were mostly straight-line patterns of the grid, squares, line maze, mosaic or radiating line type. The effects of light adaption were measured by adapting to the Aldis lamp and then suddenly interrupting the beam by the rotating episcotister. The prominent feature under these conditions was a clear-cut, complex, regular mosaic in fourfold radial symmetry made up of small elements such as may be obtained in a kaleidoscope. This differed from the mosaics obtained under other conditions which were usually line-mazes rather than mosaics, or else they were simple mosaics made up of regular squares, circles, triangles, etc. The mosaics obtained after light adaption were distinguished by the complexity of their geometrical structure and their clearness, brilliant colour and their fourfold radial symmetry. This effect was obtained fully in 6 subjects. In 2 the only specification missing was the radial symmetry. In 6 this effect was sometimes obtained or else some pattern more regular than usual. In one subject (34) another effect was seen: after dark adaption a pattern of separate squares was obtained, whereas light adaption produced a grid. In 12 subjects there was no noted difference between the two conditions. In one of these, however, a complex line maze in fourfold radial symmetry was obtained in either case. In 8 subjects no light adaption was done.

The effects of accommodation. During the experiments the eyes were normally kept at a relaxed level of accommodation. The effect of changing the degree of accommodation was tried but without any marked results. No obvious effect was noted by 14 subjects. The patterns were reported to go on the actual change but to reform as before or in the usual sequence in 10 other subjects. One subject was confident that near accommodation favoured a checker-board pattern (10/10). Another that near accommodation brought on a grid (10/11) and a spiral (1/11), whereas far accommodation brought on the same patterns in reverse ratio.

The effects of eye-movement. These again were not very exciting. The pattern merely followed the eye without changing (6 subjects); it broke up during the actual eye-movement but reformed with no significant change (11 subjects); it broke up with large eye-movements but followed small ones (3 subjects); no effect (2 subjects); the pattern went the other way to the eye-movement (2 subjects—both were certain of this effect). One subject reported that the eye-movement inhibited the dark-phase patterns and one that it brought them on. Three subjects reported that they could definitely control the direction of rotation of figures by appropriate eye-movements. This amounted to giving a little flick with the eyes. The experimenter once obtained the appearance of two large enmeshed cog-wheels slowly rotating (stimulus frequency 6 f./s.). If these were rotating so that the left wheel was going clockwise slowly raising the eyes would transform this to an anticlockwise rotation with an equivalent change in the other wheel so that they continued to rotate enmeshed as proper cogwheels. Lowering the eyes would reverse this movement and this could be controlled in this manner quite definitely for as long as this particular pattern lasted. One highly competent observer was quite sure that the dark-phase patterns did not follow eye-movement and one other subject thought that this was so.

The effect of voluntary control. This varied very much from subject to subject. It was tested by asking the subject if he could see a series of the standard patterns merely by 'willing' to see them. He had only to report definite results. The effect of volition on the movements of the pattern were also tested. Some subjects could change nothing (subjects nos. 21, 23, 24, 25, 35); one (16) could change the direction of rotation of a vortex only; five (17, 19, 20, 22, 28) had a very small amount of control: it was either transitory or doubtful or occasional; three subjects (26, 27, 29) reported a moderate degree of control (20–50 % success with the patterns) and three subjects (8, 15, 18) reported a high rate of success (above 50 %). The rest were not tested in this way. Although the series is too small to apply a valid statistical test it is probably relevant to note that 4 of the 5 subjects unable to influence their patterns in any way had very stable grade B patterns, whereas this was so in only 1 subject out of 5 in those with some degree of control. In those with a high degree of control there were two grade B subjects but these were both class 2 with the powerful alternation of two main patterns.

The effect of different flash frequencies, modes and illumination on quantity of pattern. The number of reports of pattern were counted for each of the three flash frequencies used (6, 12, 18 f./s. with both Aldis and Strobe) and significant differences were sought by Friedman's ranking method. The tabulation was done for each mode of observation (*SO*, *SR*, *OR* and *RO*) in terms of the number of reports of a nameable geometrical pattern per 100 runs. The results for white light are shown in Table 4 for both I.P. and D.P. χ^2 for frequency (using Friedmann's ranking method throughout this section) for both I.P. and D.P. is 5·6 with 2 degrees of freedom which is not quite significant at the 5 % level. This suggests only that there is the same amount of patterning at 12 and 18 f./s. and less at 6 f./s. χ^2 for modes (*SO*, *SR*, etc.) for white and monochromatic light is, for I.P., 22·1 for 3 degrees of freedom which is significant at the 0·1 % level and shows that there is more initial patterning at *OR* than in the other modes. With D.P. χ^2 is 16·6 with 2 degrees of freedom which is significant at the 0·1 % level and it confirms the previous indication that there is less patterning with single eyes than both, but how much of this is due to the effect of the dark phase patterns is doubtful (these occupy time that might have yielded a report of a bright phase pattern). χ^2 for the effect of the colour of the stimulating light on the number of patterns is 10·4 with 4 degrees of freedom which gives a p value of 0·04. This suggests that there is more patterning with white light and least with red. The effect of degree of illumination is shown in Table 5. χ^2 for the effect of different levels of illumination is 12·3 with 2 degrees of freedom which is significant at the 0·3 % level. This shows that there is progressively less patterning as the intensity of the light is decreased (as might be expected).

The effect of different flash frequencies, modes of illumination on quality of pattern. (i) Frequency. There were clear-cut differences in the form of the patterns at the three frequencies. The patterns at high frequencies are finer and made up of smaller elements than those at lower frequencies which are coarser and made up of larger elements. This is well illustrated in the further experiments to be described in part III of this paper. There were also other inconstant differences of design, but these were not enough to merit any generalizations. (ii) Modes. *SR* and *OR* are particularly associated with the parallel vertical lines pattern as shown in Table 6. (iii) Illumination. There were some qualitative changes in the form of the patterns at dim lights. In the case of the experimenter the maze/mosaic at dim levels of illumination was mainly compounded of longish zig-zag lines, whereas at higher levels short right-angled elements became more dominant and subject 3 noticed his circles were larger with dimmer light. In general, however, the patterns obtained with dim lights were merely dimmer copies of the patterns obtained with brighter lights.

Details of some specific patterns: (*a*) *Vortex.* This showed several variations. Sometimes it was described as a rotating blob and sometimes this blob showed structure such as spiralling, the lines of a brandy-ball or maypole, or 'like a swarm of bees'. (*b*) *Grid.* This was usually just a plain line grid or plaid. It could also be a lattice structure with little squares or diamonds at the intersections. It could be compound of vertical and diagonal lines, or it could have an irregular form as of a brick wall. A family of parabolas or concentric circles was sometimes superimposed on the grid (Pl. 2, fig. 12). (*c*) *Vertical lines.* These varied from subject to subject. Sometimes they were very numerous and very thin and uniform in thickness across the field. At other times they were thin at the centre and thicker at the periphery, or they could be thickest 10 degrees out from the centre. They were always however the same thickness up and down. Sometimes they were bent or had black blobs on them. They were usually in black and white but purple, green or pearly ones were also encountered (Pl. 2, fig. 10). (*d*) *Mosaic.* This could either be simple made up of small right-angled lines interlocking, or small squares and zig-zag lines (Pl. 2, fig. 9), or complex and kaleidoscopic. (*e*) *Snowflake.* This was best developed in subject 25. Here it started off characteristically as a simple snowflake and then became more and more intricate and complex and finally became disrupted by 'faults' which ripped across the pattern in all directions. The 'faults' were themselves patterned like a Greek frieze (Text-fig. 18). The final effect was one of a mass of fragmented snowflake and bits of scroll in incessant movement 'bits break up, reform and interchange'. Subject 26 reported a snowflake type of mosaic (Text-fig. 19) and subject 9 a simpler type. (*f*) *Sine-waves.* This pattern usually consisted of a series of parallel sine-waves in phase; once they occurred out of phase (Text-fig. 20) and once as the interesting double wave shown in Text-fig. 20. (*g*) *Mandala.* This was the name a few subjects gave to a line pattern in fourfold radial symmetry. It was prominent only in subject 8 (an example is shown in Text-fig. 14) where it was usually compound of straight lines, but twice of fused parabolas. (*h*) *Diffraction pattern.* This was described by only one subject (no. 11, a physicist) and was composed of a roughly radial array of squarish shapes growing larger to the periphery (Pl. 1, fig. 4). (*i*) *Chains and vortices.* This again was described only by subject 11 who was a particularly good and careful observer. The chains were strings of dots which moved in line fairly slowly from one side of the field to the other and usually rotated around a small centre or vortex on the way. They might coil down as a spiral, or several chains might wind into a vortex in a nebula effect, or a number may form a collection of parabolas.

Table 4. *Breakdown by modes, frequency and colour of total number of reports of pattern per* 100 *runs*

	I.P.			D.P.		
Strobe	12	18	6	12	18	6
SO	41	44	39	121	118	86
SR	42	39	32	73	75	57
OR	77	82	60	71	79	54
RO	41	48	34	—	—	—

Columns give f./s. Rows give modes.

	I.P.			D.P.		
Aldis	12	18	6	12	18	6
SO	56	57	40	50	60	56
SR	46	37	39	23	43	32
OR	63	50	76	28	29	32
RO	42	37	38	—	—	—

Columns give f./s. Rows give modes.

	Colour					Colour				
	R	B	G	Y	W	R	B	G	Y	W
SO	29	40	35	40	39	79	75	77	70	91
SR	12	15	24	20	18	47	42	36	41	45
OR	45	57	59	58	50	39	45	42	51	48
RO	31	36	41	37	37	—	—	—	—	—

Columns give colours. Rows give modes.

Table 5. *Breakdown by modes and level of illumination: number of reports of pattern per* 100 *runs*

	I.P.			D.P.		
	0	0·6	1·2	0	0·6	1·2
SO	58	8	29	69	59	25
SR	55	45	15	37	47	7
OR	78	76	23	39	35	12
RO	52	43	15			

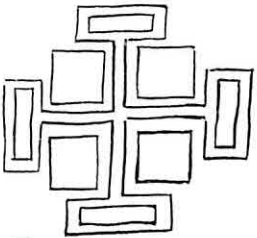
(Columns give neutral filter values, runs give modes.)

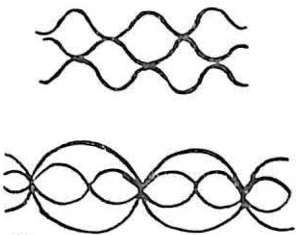
Table 6. *Number of patterns of class* 2 (*straight line*) *per* 600 *runs for each mode*

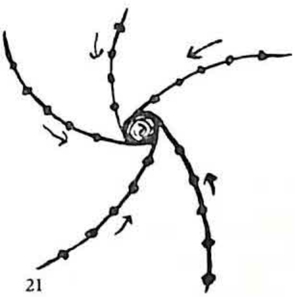
Pattern	*SO*	*SR*	*OR*	*RO*
10. Checker-board	12	3	30*	17
6. Grid	34	12	19	41
11. Vertical lines	0	58	80	2
7. Squares	30	5	9	7
Σ 10, 6, 7	76	20	58	65

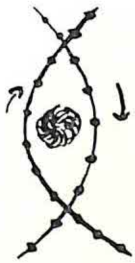
* More than 90% of the total contributed by 1 subject.

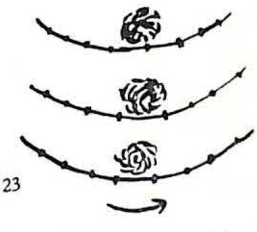
Examples of these are shown in Text-figs. 21–24. The chains appeared to 'work' the vortices as a chain does a pulley. The whole pattern was in one particular form for only a fraction of a second and then the old chains went and new ones appeared. There were usually from one to four chains in view at any one time, although there might be more at times when a whole group might flow together across the field in parallel lines. Sometimes a vortex might move across the field pulling its chains with it. The chains moved as a whole steadily across the screen. Sometimes they formed a circle and this then would develop into a vortex. A vortex might also extrude chains as it rotated, or alternate between dragging chains in and extruding them. They rarely crossed each other. Associated with this pattern was another called 'flashes' which were usually brighter than chains and moved rapidly across the screen. Sometimes they went round sharp corners and sometimes they followed the track of a chain but usually they had straighter lines and shot across the field. (*j*) *Diamond or rectangular spiral.* This was reported dominantly by one subject (no. 13). It is shown in Pl. 2, fig. 8. It varied in size from 10 degrees of visual angle in diameter to filling nearly the whole field and it could be wound either clockwise or anticlockwise. At higher speeds it was usually yellow in colour and surrounded by blue spots and at lower speeds it was blue. In the centre of the arms of the spiral white dots were sometimes to be observed streaming outwards. It was described as being of very perfect geometrical form. (*k*) *Magnetic lines of force.* These could be either arranged as around one pole of magnet, or as around two south poles, or one south and one north pole. (*l*) *Cross.* This could be either a simple line cross, bar cross or maltese cross, and either vertical or diagonal. (*m*) *Wedges.* This particular pattern was shown only by subject 14 on *OR*. There was a vertical line down the centre of the field and then a family of curved wedges filling the temporal field with some

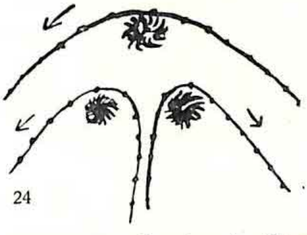

19


20


21


22


23


24

Text-figs. 19–24. Line drawings copied by the author from line drawings by the subjects made immediately after observation.

Text-fig. 19. 'Snowflake mosaic.' Subject 26. Text-fig. 20. Rare types of sine waves. Text-figs. 21–24. Types of 'chain and vortex' pattern. Subject 11.

scrolls in the periphery. This was described as being just like an appearance of turbulence behind an object in a wind tunnel. The nasal side of the field was either clear or had only one wedge in it (Pl. 2, fig. 11).

The other patterns did not contain features of note.

Relation of curved to straight lines. These only very rarely formed connected elements of a pattern. The only recorded reports were one of a square with rounded corners and the 'thin blue line' pattern of subject 1 (Pl. 2, fig. 7). However, a straight-line pattern and a curved-lined pattern could be superimposed (Pl. 2, fig. 12) or one could be in the centre and the other in the periphery of the field (Pl. 1, fig. 5).

Other types of pattern never seen. Patterns in threefold radial symmetry were extremely rare and in fivefold symmetry almost as rare.

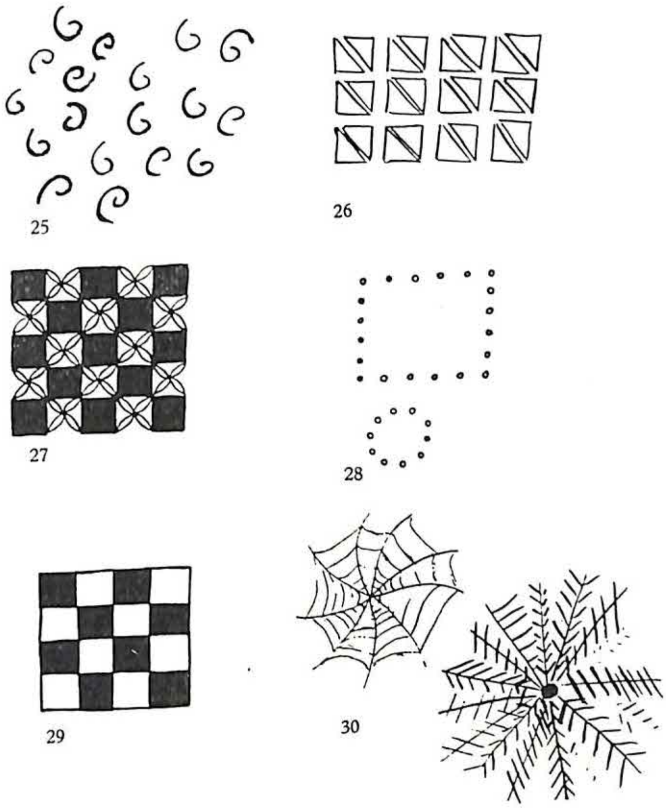
The colours of the patterns. These again varied greatly from individual to individual. In a few cases only blacks, whites and greys were observed; others reported various pastel shades, others again brilliant and saturated colours. Some subjects reported only a few colours, others very many, but no constant regularities or correlations were observed. All the colours of the spectrum and intermediate colours were obtained.

(4) *Experimental precautions*

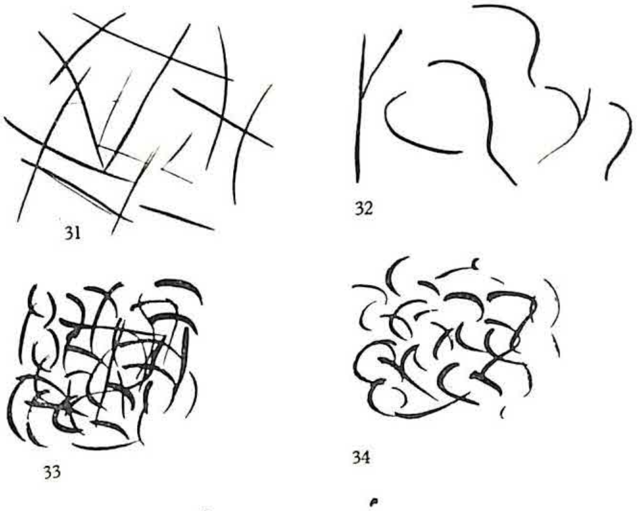
As the patterns are, to a greater or lesser extent, under some voluntary control, great precautions were taken during these experiments to ask no leading questions. All questions were asked in terms such as 'What pattern do you see?' In some instances, specific patterns were inquired after: in which case a note to this effect was made in the experimental protocols and the effect of possible suggestion was borne in mind. Other difficulties that had to be taken into consideration were as follows. Subjects were not always able to give a full description of all changes going on in the field. Either too many events might be going on at the same time, or some of the events might not be easily describable. Furthermore, there is a limit to the detail in which such patterns can be described. For example, certainty might be expressed that a star or concentric circles were visible but it would only be possible to give the size or angularity of the elements within broad limits. At times the number of elements could be given with confidence, but at other times only within limits. However, the significant generalizations here reported are based only on major features of the patterns that were reported with certainty.

(5) *After-images*

After intermittent stimulation at all the flash frequencies and intensities of illumination used, after-images, both geometrical and non-geometrical, were reported. The geometrical patterns were in general both qualitatively and quantitatively similar to the bright-phase geometrical patterns except for a relative scarcity of vortices, crosses and vertical parallel lines in the case of the after-images. There were considerably fewer geometrical patterns with *RS* than with *OS*. However, many more non-geometrical patterns of a curious type described as random thin curly or squirly lines often in movement, or 'spaghetti', 'filigree', 'cottonwool', or 'worms', or as a swirling effect described as like smoke or 'wood smoke' occurred with *RS* then with *OS*. Table 7 expresses this in detail. There were no very great qualitative differences between the types of geometrical pattern seen with *OS* and *RS*, nor between those seen with white and monochromatic light. There was some difference at the different frequencies. At 12 f./s. there were 41·2 patterns reported per 100 runs, at 18 and 6 f./s. there were 34·7 and 27·1 patterns reported per 100 runs respectively.



Text-figs. 25–30. Line drawings copied by the author from line drawings by the subjects made immediately after observation. Types of binocular after-images. Subjects 26, 18, 26, 14, 1, 17, respectively. Colours: 26, red and green in pairs on black; 27, crimson, yellow and black; 29, various—black and green, purple and green, black and white, brown and green, black, white and yellow; 30, green on black.



Text-figs. 31–36. Types of uniocular after-images of the 'spaghetti' kind. Same subjects as Text-figs. 25–30, except for Text-fig. 35 (subject 15) and Text-fig. 33, which corresponds to the verbal description of subject 2 and not a drawing.

The after-image was in most instances a pattern different from the immediately preceding bright-phase pattern. Very occasionally in a few subjects and commonly in only one subject the last bright-phase pattern carried over briefly as a positive after-image. The after-images may develop and change for up to 30 sec. Examples are given in Text-figs. 25–36.

Subjects may be classified into five classes on a basis of their after-images.

Class A.I. (1). *One dominant pattern*

Class A.I. (1*a*). Many small squares, circles, diamonds, triangles, and/or stars all over the field

Subject	No. of different patterns	'Spaghetti' type	Dark phase	Bright phase Grade	Bright phase Class
12	6	0	0	A	1
5	4	±	+	B	1
26	5	+	+	A	1
25	2	+	+	B	7

Class A.I. (1*b*). 'Stained glass window' effects

18	1	+	+	A	1

Class A.I. (1*c*). One dominant pattern other than *a* or b, given in parentheses

1	1 (checker-board)	0	0	B	4
7	3 (ellipses)	0	+	B	unclass.
17	4 (complex herring-bone)	+	+	A	5
13	1 (checker-board)	±	+	B	4
28	5 (grid)	±	+	A	1
22	4 (grid)	+	0	C	1
3	3 (regular pattern of dots)	0	0	B	1
14	3 (regular pattern of dots)	±	+	B	3

Class A.I. (2). *Frequent but different patterns*

29	5	0	0	C	6
30	8	0	0	A	1
2	6	+	+	B	2

Class A.I. (3). *Infrequent patterns, none dominant*

21	4	0	+	A	1
34	2	+	0	A	1
31	2	0	0	A	5
27	2	0	0	B	2
15	1	+	+	A	1
6	2	+	+	A	unclass.
32	3	±	0	A	1
8	5	0	0	B	2
16	4	+	+	A	1

Class A.I. (4). *No patterns*

11	0	0	+	B	1
19	0	0	0	A	6
24	0	0	+	B	6
20	0	0	0	B	1
9	0	+	+	B	1
4	0	0	0	C	1
10	0	0	0	B	1

Class A.I. (5). *Positive after-image of the bright phase patterns only*

23	–	0	0	B	1

It will be noticed that there is no obvious correlation between the class and grade of the bright phase patterns and the type of after-image obtained in each case. One significant fact was, however, that the 'spaghetti' type after-image occurs more frequently in these subjects showing dark-phase patterns than in those not showing these patterns. Table 7 expresses this quantitatively ($\chi^2 = 10{\cdot}1$, 1 degree of freedom: $p < 0{\cdot}01$). Each + in the tabulations is scored 1 and each ± is scored ½. ± scores were awarded once where the 'spaghetti'-type image was only infrequent, three times where it was atypical, being composed of swirling smoke or 'wood smoke' rather than random thin lines, and once where it was atypical, being composed of 'seething sand or grains'.

Table 7. *Correlation of 'spaghetti'-type after-images and the dark-phase patterns*

'Spaghetti'-type	Dark-phase patterns: Present	Dark-phase patterns: Absent
Present	11	2·5
Absent	4	13

Many of the geometrical after-images were described as being more clear-cut and definite than the rather fuzzy and continually agitated bright-phase patterns. The 'stained glass window' effect is well described by subject 18 where it occurred most often, although a similar pattern was described at times by other subjects. He described it as follows: 'triangles in red, green and some white like a stained glass window separated as if by lead and in a regular pattern—very rich colours'.

IV. SUMMARY OF POSITIVE FINDINGS

This paper gives an account of the bright-phase patterns that are visible under conditions of intermittent stimulation of the retina. The following main facts have emerged. (i) The range and type of pattern observed and their behaviour is characteristic for each individual person; yet there are sufficient similarities between people to enable classification to be carried out. I have distinguished seven classes and three grades. (ii) The types of pattern are also classifiable and fall rather clearly into radial patterns, straight-line patterns on a square base, herring-bone and honeycomb types, curved-line patterns and complex labyrinth/maze/mosaics. (iii) A variety of movements of the patterns is reported and each pattern has a specific range of movement. (iv) There are characteristic differences between uniocular and binocular stimulation; between the centre and periphery of the field; between high and dim levels of illumination; between fast and slow rates of flicker; between dark- and light-adapted states of the retina, and between various other different parameters. (v) Characteristic after-images may also be obtained similarily characteristic for the individual and classifiable between individuals. These also respond to parameter change. (vi) Perhaps the most significant finding has been that the patterns are only very, very rarely composed of straight *and* curved lines as connected elements of the pattern: they are either straight-line patterns or curved-line patterns.

Part III of this paper will contain an account of further experiments and discussion.

I am most grateful to Prof. O. L. Zangwill for his continued interest in and support of these experiments; to Prof. M. D. Vernon and Dr W. A. H. Rushton for their invaluable criticisms of an earlier draft of this paper that have enabled me to avoid some serious mistakes; to Sir Russell Brain for his aid in the initiation of this work and to Prof. Heinrich Klüver for the benefit of many fascinating discussions.

REFERENCES

AUBERT, H. (1865). *Physiologie der Netzhaut*, pp. 355, 377–80. Breslau.

BEURLE, R. L. (1956). Properties of a mass of cells capable of regenerating pulses. *Phil. Trans.* B, **240**, 55–94.

BLUM, A. C. (1956). Photic stimulation, imagery and alpha rhythm. *J. Ment. Sci.* **102**, 160–7.

BREWSTER, D. (1861). On certain affections of the retina. *Phil. Mag.* **21** (4th series), 20–4.

BROWN, C. R. & GEBHARD, J. W. (1948). Visual field articulation in the absence of spatial stimulus gradients. *J. Exp. Psychol.* **38**, 188–200.

COSTA, A. M. (1953). L'effetto geometrico-cromatico nella stimolazione intermittente della retina ad occhi chiuse. *Arch. Psicol. Neur. Psich.* **14**, 632–5.

ELRIDGE–GREEN, F. W. (1925). *Physiology of Vision*. London.

EXNER, S. (1870). Bemerkungen über intermittierende Netzhautreizung. *Pflüg. arch. ges. Physiol.* **3**, 214–40.

FECHNER, G. T. (1838). Ueber eine Scheibe zur Erzeugung subjectiver Farben. *Pogg. Ann. Physik. Chemie*, **45**, 227–32.

GEBHARD, J. W. (1943). Chromatic phenomena produced by intermittent stimulation of the retina. *J. Exp. Psychol.* **33**, 387–406.

GREY WALTER, W. (1950). In *Perspectives in Neuropsychiatry* (D. Richter, ed.). London.

HELMHOLTZ, H. VON (1924). *Physiological Optics*. Optical Society of America.

MUNDY–CASTLE, A. C. (1953). Electrical responses of the brain in relation to behaviour. *Brit. J. Psychol.* **44**, 318–29.

PAULI, R. & WENZL, A. (1924). Über Farbenempfindungen bei intermittierendem farblosen Lichte. *Arch. ges. Psychol* **48**, 470–84.

PURKINJE, J. E. (1823). *Beobachtungen zu Versuche für Physiologie der Sinne. Beiträge zur Kenntnis des Sehens in subjectiver Hinsicht.* Bd. I. Prague.

SMYTHIES, J. R. (1959). The stroboscopic patterns. I. The dark phase. *Brit. J. Psychol.* **50**, 106–16.

UTTLEY, A. M. (1954). The classification of signals in the nervous system. *EEG Clin. Neurophysiol.* **6**, 479–94.

EXPLANATION OF PLATES

PLATE 1. Photographs of pastel drawings made by the subjects immediately after observation

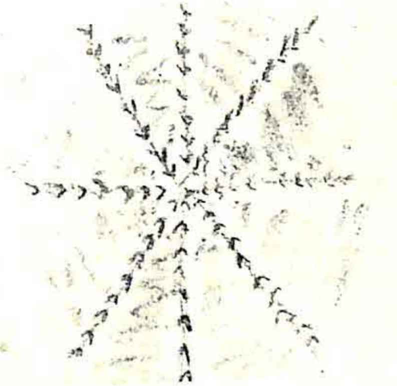
Fig. 1. Star figure made of herring-bone. I.P., subject 3. Colours: black and white.
Fig. 2. Banded circles figure. D.P., subject 3. Colours: black and white.
Fig. 3. Star figure. I.P., subject 9. Colours: yellow, red and black on grey.
Fig. 4. Diffraction pattern. Subject 11. Colours: brown on yellow.
Fig. 5. Spiral with peripheral mosaic. D.P., from subject 8. Colours: blue on yellow.
Fig. 6. Scallop pattern. D.P., subject 14. Colours: red and blue on yellow.

PLATE 2. As Pl. 1

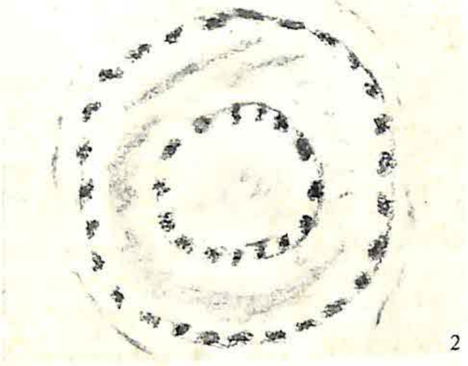
Fig. 7. Thin blue line figure. I.P., subject 1. Colours: blue on white.
Fig. 8. Diamond spiral. D.P., subject 13. Colours: spiral, yellow: dots, blue on very pale grey.
Fig. 9. Maze/mosaic. D.P., subject 35. Colours: blue, yellow (at centre) black, olive-green and grey on pale grey.
Fig. 10. Vertical line pattern. I.P., with *OR*, subject 32. Colours: black and purple on white.
Fig. 11. Wedge pattern. I.P. with *OR*, subject 14, Colours: black on white.
Fig. 12. A family of parabolas superimposed on a grid from a subject outside the series. Colours: parabolas, yellow and black: grid, brown and purple, on pale grey.

(*Manuscript received* 30 *January* 1958)

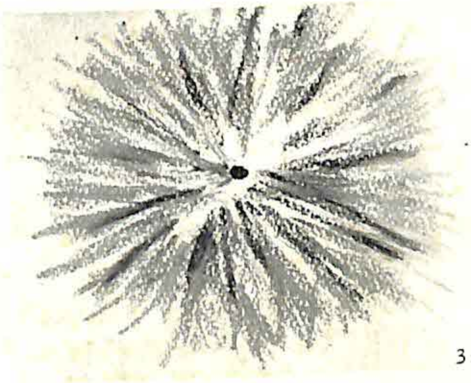
Plate 1



1

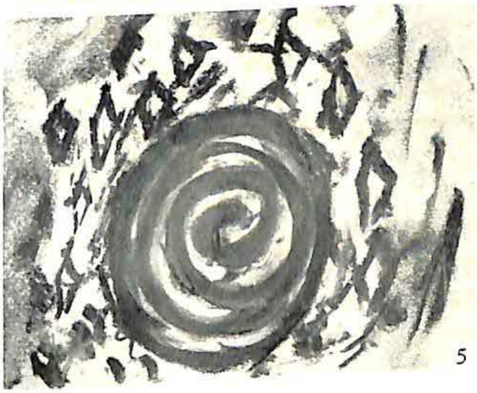


2

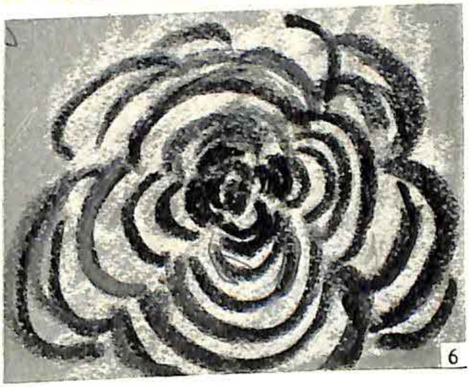


3

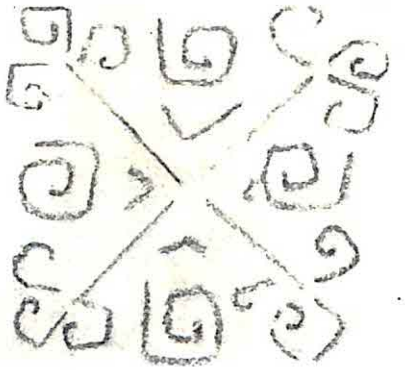
4



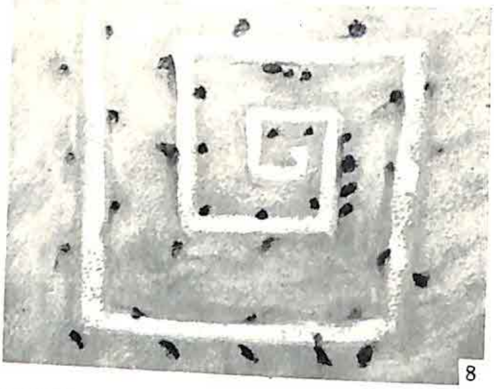
5



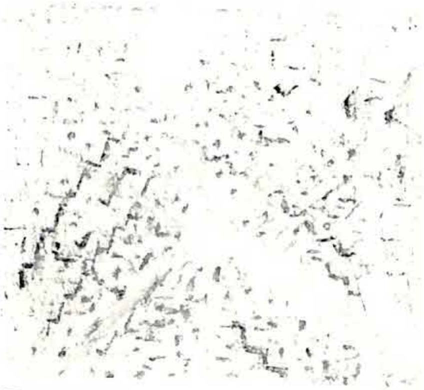
6



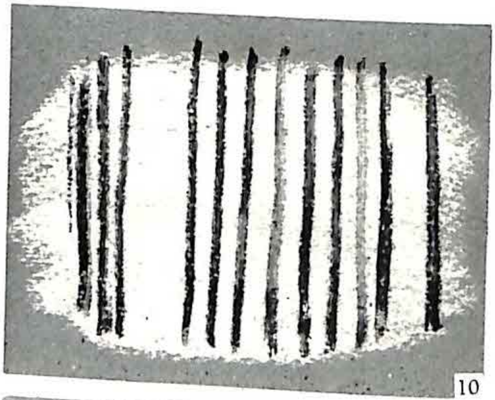
7



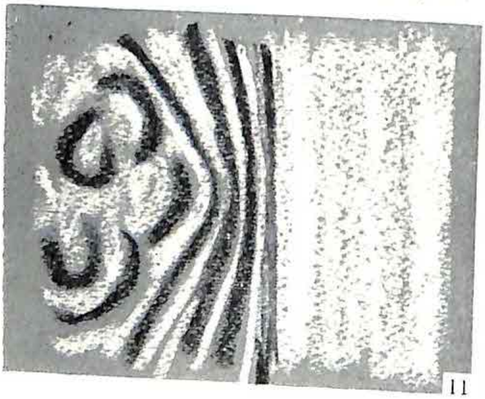
8



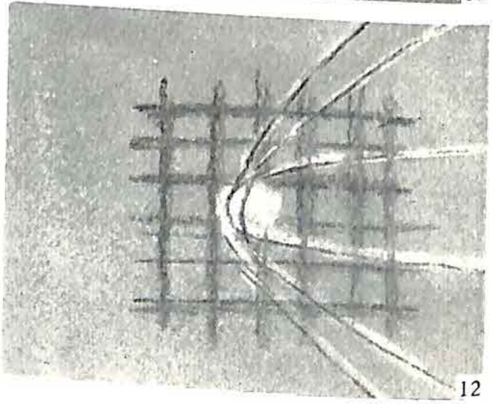
9



10



11



12

EXTRAVERSION—INTROVERSION AND IMPROVEMENT IN AN AUDITORY VIGILANCE TASK*

By PAUL BAKAN

Michigan State University†

The study deals with the effect of a secondary signal-detection task on performance in a primary signal-detection task in an auditory vigilance situation. Forty subjects were tested under two conditions. In one condition they listened to an 80-min. recording of a sequence of digits in order to detect the occurrence of primary signals defined as *three successive odd digits which are all different*. In another condition they listened for primary signals as described above, but in addition they also had to detect secondary signals defined as the occurrence of *the digit* 6. The primary signals occurred ten times every 16 min.; there were about ten times as many secondary signals (digit 6) as there were primary signals (odd–odd–odd).

The over-all detection of primary signals was better when subjects were listening for both primary and secondary signals than when they listened for primary signals alone. Extraverts benefited more by the secondary task than introverts. No significant relationship was found between either neuroticism or intelligence and vigilance performance. A theoretical account of the events in a vigil is proposed and the difference between introverts and extraverts is considered in terms of this account.

I. Introduction

Vigilance has been defined by Mackworth (1957) as a state of readiness to detect and respond to certain specified small charges, occurring at random time intervals in the external environment. Experimentally the level of vigilance is inferred from the performance of the subject by considering the number of signals which are not responded to (Mackworth, 1950) or the intensity to which signals have to be raised in order to be responded to (Bakan, 1955; Elliott, 1957).

Much of the experimental work with vigilance tasks has been concerned with environmental variables, either in the work situation, or in the general surroundings. Experiments have been designed to evaluate the effects of task duration (Bakan, 1955; Mackworth, 1950; Wittenburg, Ross & Andrews, 1956), signal frequency (Deese, 1955), signal intensity (Adams, 1956; Mackworth, 1950), signal duration (Adams, 1956), inter-signal interval (Baker, 1956), noise (Broadbent, 1953; Jerison & Wing, 1957; Jerison & Wallis, 1957; Loeb & Jeantheau, 1958), temperature (Pepler, 1953), and isolation (Fraser, 1953). In some instances the state of the subject has been varied, as in studies on the effect of drugs (Mackworth, 1950), sleep deprivation (Wilkinson, 1957), rest periods (Mackworth, 1950), and knowledge of results (Mackworth, 1950).

Large individual differences in both over-all performance and in the course of performance over time are characteristic of performance in vigilance tasks. Though most investigators have reported wide individual differences in performance on vigilance tasks, there has been very little research on the correlates of these differences, though some workers have suggested the possible importance of personality factors in accounting for them (Deese, 1955; Fraser, 1953).

---

* An earlier draft of this paper has appeared in unpublished form as M.R.C. Applied Psychology Unit Report No. 311 (1957). The author wishes to express his thanks to the members of the Applied Psychology Research Unit for their critical interest in the research and especially to Sir Frederic Bartlett, Donald Broadbent, and Dr N. H. Mackworth.

† The research reported in this paper was carried out in 1956–57 while the author was a post-doctoral Fellow of the National Science Foundation at the M.R.C. Applied Psychology Research Unit, Cambridge. The work was done under the general direction of Dr N. H. Mackworth.

The present study is concerned with the evaluation of the effect of an environmental task variable on performance in an auditory vigilance task, and the relationship between personality variables and vigilance performance. The task variable consists of the addition of a secondary task to a primary vigilance task resulting in a large increase in the total number of relevant stimuli to be responded to. The problem is to determine whether the introduction of a secondary task results in an increase in signals detected on the primary task. It has been shown that vigilance performance improves as the number of relevant signals increases where all of the signals are of the same kind (Deese, 1955). In most practical situations, however, it is not feasible to increase artificially the number of signals when 'real' and 'artificial' signals are of the same kind and when responses to both are identical. Therefore, in the present experiment a secondary task was added to a primary vigilance task, resulting in an increase in the *total* number of relevant signals coming from an auditory display. However, the primary and secondary signals, though coming from the same display, were different and required different responses to be made. The hypothesis under consideration was that subjects would detect more primary task signals when listening for both primary and secondary task signals than when listening for primary task signals alone.

In addition to measures of performance on the vigilance task, information was available for each subject on two personality variables, neuroticism and extraversion–introversion. Analyses were carried out to test the hypothesis that personality variables are related to individual differences found in performance on vigilance tasks.

II. METHOD

(i) *Stimulus material*

The stimulus material consisted of a tape-recording consisting of digits presented at the rate of one per second. The randomness of these digits was somewhat restricted but they gave the general impression of a random series. Subjects were required to listen for a signal which was defined as *three successive odd digits which are all different*. The frequency of occurrence of such signals was ten every 16 min. The entire tape consisted of five 16-min. periods with no break between periods. The 16 min. subperiods were equivalent with respect to number of signals and their distribution in time. Within each subperiod the time in seconds between successive signals was as follows: 71, 154, 25, 185, 110, 104, 46, 15, 143. The time between the last signals of one subperiod and the first signal of the next subperiod was 146 sec.

(ii) *Primary and secondary task*

The primary task consisted of listening for signals (primary) as described above (odd–odd–odd), and writing down the actual digits constituting the signal whenever it was detected. The secondary task consisted of listening for the digit 6, the secondary signal, and pressing a button whenever a 6 was detected. The secondary signal also appeared irregularly but with a far greater frequency than the primary signal. The digit 6 appeared on the average of once every 10 sec., as contrasted with an average of about once every 1·6 min. for primary signals. Subjects were told that of the two tasks the primary one was the more important in contributing to their score, but that they should get as many of both kinds of signals as possible. The analyses in the present paper are based on detection of primary signals.

(iii) *Subjects*

Subjects were Royal Naval ratings. There were forty subjects in this experiment and each subject served in two conditions. On one day a subject would listen for primary signals alone (0–0–0), and on another day he would listen for primary and secondary signals (0–0–0 + 6). Half the subjects did the primary task on day 1 and the primary plus secondary task on day 2, and the other half did the tasks in the reverse conditions. Experimental sessions were usually about a week apart.

(iv) *Practice*

After it was clear that the subject understood the instructions he listened to a practice tape of 10·5 min. which contained six primary signals (0–0–0) and ninety-seven secondary signals. The practice tape was used both days for each subject and the practice corresponded to the experimental condition for that day. During practice there was knowledge of results in the form of the experimenter calling out signals about 3 sec. after their passage. Immediately after the practice period there was a 10-min. break during which the subject left the testing room.

(v) *Test conditions*

During the 80-min. test period which followed the break, subjects were left alone in a cubicle in the room with the tape-recorder going. They were instructed not to leave until they were told the experiment was over. Wrist-watches were taken away for the duration of the test session. Subjects were not permitted to smoke during the test period.

(vi) *Other available data*

Measures of intelligence, neuroticism, and introversion–extraversion were available for each subject. The intelligence measure available was the A.H. 4 Intelligence Test, and the personality measures came from the Heron Personality Inventory.

III. RESULTS

(i) *Over-all effect of secondary task on primary task performance*

The effect of the secondary task was to improve the over-all performance on the primary task. Of the forty subjects there were thirty-nine who did not get perfect scores with the primary task alone and thus could show improvement. Of these thirty-nine there were twenty-three who improved with the addition of the secondary task, eleven who did worse with the secondary task and five whose performance was the same under both conditions. The improvement with the secondary task is significant by a sign-test at better than the 5% (1-tail) level. Over the 80-min. test-period the mean number of failures in detecting a primary signal when subjects were listening for primary signals alone was 9·95; when listening for both primary and secondary signals together the mean number of detection failures was 7·90. This improvement was independent of the order of administration of experimental conditions.

(ii) *Subject stability between conditions*

In order to determine whether over-all performance under one condition was related to over-all performance on the other condition, a correlation was computed between signal omissions on the primary task alone (0–0–0) and omissions on the primary and secondary tasks (0–0–0 + 6). The correlation coefficient (Pearson) between omission scores on these two tasks was 0·71, which is significant at the 1% level. This means that despite the improvement taking place with the secondary task, the relative standing of subjects was not changing very much. This tends to suggest that the performance of subjects on vigilance tasks may be relatively stable.

(iii) *Introversion–extraversion and vigilance*

Measures of introversion ranging from 0–9 were available for all subjects. A split was made as close to the median as possible to divide the subjects into two subgroups, extraverts ($N = 19$) and introverts ($N = 21$). All subjects with scores of three or less were classed as extraverts and all subjects with scores of four or more were classed as introverts. There was no significant difference either in neuroticism or intelligence between these groups. It was found that the extraverts were more likely to improve with the secondary task than the introverts. Of the nineteen extraverts there were fourteen who improved with the secondary task, whereas only nine of the twenty-one introverts improved with the secondary task. This relationship between introversion–extraversion and over-all improvement with the secondary task was significant ($\chi^2 = 3{\cdot}97$; $p = 0{\cdot}05$).

Analysis of the data by 16-minute sub-periods results in further information on the nature of this relationship (see Fig. 1). The improvement of the extraverts with the secondary signal seems to be related to the fact that in the first two subperiods (32 min.) their performance in detecting primary signals (with no secondary signals) is worse than that of the introverts (1-tail, U-test; $p = 0{\cdot}03$). With the introduction of the secondary task there is no significant difference between introverts and extraverts in the first two subperiods. The main effect of the secondary task is to improve performance of the extraverts in the first two subperiods. The result is that the extraverts with the secondary task behave like the introverts without it with respect to the detection of primary signals. For periods 1 and 2 the extraverts do significantly better with the secondary task than without it (sign-test, 1-tail; $p = 0{\cdot}001$ for period 1 and $p = 0{\cdot}029$ for period 2). The difference is not significant for the remaining periods (3, 4 and 5), though there are small differences in the direction of better performance with the secondary task.

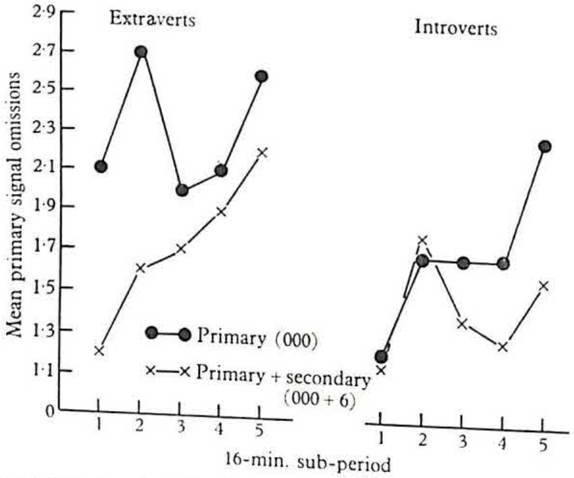


Fig. 1. Relationship between mean number of primary signals omitted and time in the experimental situation under primary signal condition and primary plus secondary signal condition for subjects classified as extravert and introvert.

The effect of the secondary task on the performance of intraverts is different. In the first two periods there is no appreciable difference in primary signal detection as a function of the secondary task. But for periods 3, 4 and 5 there are differences in the direction of improved performance with the secondary task, with the difference between conditions reaching significance in the fifth period (sign-test, 1-tail; $p = 0{\cdot}006$).

The evidence presented so far indicates that there is a relationship between introversion–extraversion and improvement with a secondary task which serves to increase considerably the total number of signals to be detected. Extraverts benefit from this secondary task during the first 32 min. of the session, whereas introverts do not. This difference seems to be due to the fact that without the extra task the extraverts do significantly worse than the introverts in the first two periods, but with the extra task they improve significantly. A correlational analysis of this relationship would indicate whether the effect of the extra task is related to the *degree* of introversion–extraversion. The correlation between score on introversion-extraversion and algebraic differences between primary signal omissions in the two experimental conditions for the first two periods is $-0{\cdot}38$ (significant at 0·05 level).

(iv) *Neuroticism, intelligence, and vigilance*

A number of analyses were carried out to determine the nature of the relationship between neuroticism scores from the Heron Inventory and the scores on the vigilance tasks. No significant relationships were found in the sample studied. Relationships between intelligence, as measured by the A-H-4 test, and vigilance scores were also insignificant in the sample studied.

IV. DISCUSSION

The main findings of this experiment were (*a*) that the addition of a secondary task, which markedly increases the number of signals to be responded to, results in improved signal detection performance in a primary task, and (*b*) that extraverts, who do not do as well with the primary task alone, show greater improvement than introverts with the addition of the secondary task. These results need to be considered in the context of the particular experimental arrangements used.

In considering the effect of the secondary task on primary task performance it should be noted that the secondary task used in this experiment was easy, was compatible with the primary task, required detection of signals from the same display as the primary task, and was presented to the subjects as the less important of the two tasks. It is likely that there are many possible relationships between primary and secondary tasks which would result in findings quite different from those of this experiment.

The relationship found between introversion–extraversion and performance in the vigilance task should also be considered in the context of the details of the experiment. It should be noted that subjects were tested individually under conditions of isolation. It is difficult, therefore, to make a generalization about the relationships between introversion–extraversion and vigilance performance because of the possibility that social isolation may have enhanced the differences between introverts and extraverts. Further studies need to be carried out to analyse the interaction between personality variables and situational variables in vigilance situations.

One of the main features of a vigilance task is the freedom of the subject in determining the extent and distribution of his attentive activity with respect to the relevant stimulus material. At the beginning the subject may be instructed to pay attention and to detect as many signals as he can, but once the vigil begins he is usually free to make his own decisions about how much to listen to, or look at, the relevant stimulus display for the appropriate signals. He is also free to attend to irrelevant stimuli, either external or internal (e.g. day-dreams) and this may be incompatible with efficient detection of relevant signals. In the present experiment the freedom of the subject was probably enhanced by the absence of both the experimenter and knowledge of results during the test session.

In a situation where the subject has so much freedom, we may expect to find rather wide variation among individuals, and such variation has certainly been characteristic of data from vigilance experiments. The discovery of a relationship between a personality variable, introversion–extraversion, and performance in a vigilance task helps to account for some of this variation between subjects.

It seems that extraverts are more likely to listen when there are more signals to listen for, whereas the listening behaviour of introverts is less influenced by frequency, at least in the early part of the vigil. The data also suggest that the higher signal rate maintains the level of performance of the introverts for a longer time and reduces their tendency to show a decrement.

Before considering further the reasons for the difference between introvert and extravert performance in the vigilance task it would be helpful to try to construct a general account of what may happen in a vigilance situation. It is assumed that when subjects are instructed to listen for certain types of signals, the occurrence of these signals in the stimulus display serves to reinforce the attentive behaviour which makes the detection of signals possible. Holland has shown the reinforcing effect of signals in a visual task (Holland, 1957). If the attentive behaviour is not sufficiently reinforced by the appearance of signals, then it is possible for the behaviour to undergo extinction. To put it in other words, when a subject listens and he frequently hears what he is listening for, he is more likely to continue listening than when he listens and only rarely hears what he is listening for. The same thing would be true for a visual task if we substitute 'looks' in place of 'listens'. In a vigilance task with a low signal frequency the lack of reinforcement in the form of signals may induce the subject to stop listening.

Several consequences may follow when a subject stops listening. Though he may stop listening for signals he doesn't necessarily stop hearing the stimulus material. (Likewise he may stop looking for signals, but may still see the display where the signals may appear.) When a subject is hearing the display but not actively listening for signals, the display becomes dedifferentiated as far as the subject is concerned since he is not making discriminations within the display. To be specific, in the vigilance task used in the present experiment, when a subject stops listening for signals he stops making the discriminations required by the task of listening for sequences of three successive odd digits, which are all different. These discriminations involve successiveness of digits, oddness–evenness of digits, identity of digits, and memory for previous digits while listening for current digits. Making these discriminations in the stimulus material has the effect of differentiating the material. When the discriminations are not made, i.e. when the subject stops listening for signals, he hears instead undifferentiated sound, or a sequence of undifferentiated numbers. The effect of this dedifferentiation of the sound is to produce monotony in the afferent stimulation, which is conducive to drowsiness or sleep. In fact, reports of drowsiness are very common from subjects in vigilance tasks (Bakan, 1955, 1957). The consequence of this state of drowsiness may be sleep itself, or an attempt on the part of the subject to redifferentiate his stimulus world by means other than listening for signals. The latter course is usually chosen by human subjects.

The means used by subjects to redifferentiate their environment may include attending to irrelevant (from the point of view of the experimenter) stimuli in the external environment, or attending to internal stimulation in the form of daydreams or irrelevant mind-wandering. These irrelevant activities are usually incompatible with efficient detection of signals. The point is, that when a subject stops doing what he should be doing he may initiate a chain of events which further decreases the probability that he will again do what he should be doing. The sequence consists of the following steps.

1. Initial inattentiveness due to low signal frequency, or low reinforcement frequency;
2. Dedifferentiation of stimulus material in the relevant display;
3. Drowsiness due to undifferentiated stimulation;
4. Redifferentiation of stimulus environment by attention to irrelevant stimuli and mind-wandering;
5. Inefficient signal-detection behaviour.

With this general construction of the events in a vigilance task we may now turn to the observed difference between extraverts and introverts in the early part of a vigil, and the improvement of performance of extraverts with the introduction of the secondary task.

It appears that the difference between extraverts and introverts is in the first step of the sequence of events listed above. It seems that in the early part of the vigil the listening behaviour of the extravert is influenced more by the frequency of occurrence of signals. When the extravert listens for signals and doesn't hear any he tends to stop listening (step 1). The introvert persists in listening for signals even though they occur infrequently. This difference between introvert behaviour and extravert behaviour can be thought of in terms of the *extinction* of the responses involved in listening for signals. If the occurrence of signals reinforces these listening responses, and if signals occur infrequently, then there is not much reinforcement of the listening response, and the result is a tendency for these responses to be extinguished. This means that when signals are infrequent the subject is less likely to listen for them. It may be that extinction of listening responses takes place earlier for extraverts than for introverts. A relationship between extinction and introversion–extraversion has been suggested by Eysenck (1955) and Franks (1957).

The addition of the secondary task, which involves about a tenfold increase in signals to be detected, should serve to increase the reinforcement of listening responses and prevent the early extinction of these responses. This would explain the early improvement of the extraverts with the secondary task.

The secondary task also seems to produce some improvement in the introverts, but in this case the improvement comes in the latter part of the vigil. It can be assumed that eventually there is extinction of listening responses for the introverts, but this tends to occur later in the vigil and produces a decrement in performance. The secondary task seems to help prevent this decrement.

V. References

Adams, J. A. (1956). Vigilance in the detection of low-intensity signals. *J. Exp. Psychol.* **52**, 204–8.

Bakan, P. (1955). Discrimination decrement as a function of time in a prolonged vigil. *J. Exp. Psychol.* **50**, 387–90.

Bakan, P. (1957). Vigilance: Discussion III. *The Advancement of Science*, **53**, 410.

Baker, C. H. (1956). Biasing attention to visual displays during a vigilance task. Medical Research Council Applied Psychology Research Unit 227/56, RNP 58/876.

Broadbent, D. E. (1953). Classical conditioning and human watch-keeping. *Psychol. Rev.* **60**, 331–9.

Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.* **44**, 295–303.

Deese, J. (1955). Some problems in the theory of vigilance. *Psychol. Rev.* **62**, 359–68.

Elliott, E. (1957). Auditory vigilance tasks. *The Advancement of Science*, **53**, 393–9.

Eysenck, H. J. (1955). A dynamic theory of anxiety and hysteria. *J. Ment. Sci.* **101**, 28–51.

Franks, C. M. (1957). Personality factors and the rate of conditioning. *Brit. J. Psychol.* **48**, 119–26.

Fraser, D. C. (1953). The relation of an environmental variable to performance in a prolonged visual task. *Quart. J. Exp. Psychol.* **5**, 31–2.

Holland, J. G. (1957). Technique for behavioural analysis of human observing. *Science*, **125**, 348–50.

Jerison, H. J. & Wing, S. (1957). Effects of noise and fatigue on a complex vigilance task. Technical Report 57–14, WADC, U.S.A.F.

Jerison, H. J. & Wallis, R. A. (1957). Experiments on vigilance: Performance on a simple vigilance task in noise and quiet. Technical Report 57–318, WADC, U.S.A.F.

Loeb, M. & Jeantheau, G. (1958). The influence of noxious environmental stimuli on vigilance. *J. appl. Psychol.* **42**, 47–9.

Mackworth, N. H. (1950). *Researches on the Measurement of Human Performance.* London: H.M.S.O. Medical Res. Council Spec. Rep. Ser. No. 268.

Mackworth, N. H. (1957). Some factors affecting vigilance. *The Advancement of Science*, **53**, 389–93.

Pepler, R. D. (1953). The effect of climatic factors on the performance of skilled tasks by young European men living in the tropics. 4. A task of prolonged visual vigilance. M.R.C. App. Psych. Unit Rep. No. 156.

Whittenburg, J. A., Ross, S. & Andrews, T. G. (1956). Sustained perceptual efficiency as measured by the Mackworth 'clock' test. *Percept. Mot. Skills*, **6**, 109–16.

Wilkinson, R. T. (1957). Effects of lack of sleep. Interim Report (OES 287, FPRC 961·3) Applied Psychology Research Unit, Medical Research Council, 15 Chaucer Rd, Cambridge, England.

(*Manuscript received* 14 *July* 1958)

LABILITY IN SCHIZOPHRENIA

By N. O'CONNOR and J. P. DAS

Medical Research Council, Social Psychiatry Research Unit,
Institute of Psychiatry, London

Paranoid and non-paranoid schizophrenics and a control group of normal males were matched for age and compared on three measures of lability. Lability was first measured in terms of the ratio of trials to condition and to extinguish or the ratio of trials to extinguish and recondition. Secondly, it was measured in terms of latencies of response before and after extinction, and thirdly in terms of response pressure before and after extinction. The measures of lability were obtained with each of three groups of subjects under two stimulus conditions, one in which the stimulus was a light and another in which it was a spoken word.

Considerable differences in ease of extinction were apparent between the two types of signal. With the visual (non-verbal) signal, extinction was achieved much more readily than with the verbal signal, irrespective of diagnostic group. This result seemed contrary to expectation on the basis of Pavlovian theory. It was incidentally observed that the variability of latency scores was greater in schizophrenics than in normals as expected, but that the variability of pressure scores was less.

I. INTRODUCTION

Pavlov (1938, 1941) discussed three concepts when considering the role of excitation and inhibition in the nervous system. The first two of these appear to concern the nervous system itself and the third the reaction of the organism to signals from outside. The first is relative excitability or inhibitibility. The second is the tendency of the organism to move from one to the other state as judged by ease or difficulty in conditioning or extinguishing a response: this facility of exchange was called lability or mobility. The third concept is hard to distinguish from the other two and, strictly speaking, concerns another type of relationship. For this reason it is not further discussed in this article.

Pavlov thought that a normal organism would show a balance between excitation and inhibition which enabled conditioning and extinction to take place equally readily. However, in pathological conditions of excitability or inhibition, lability will be interfered with, making extinction more difficult in the former case and conditioning or reconditioning more difficult in the latter.

In his discussions of schizophrenia, Pavlov (1938, 1941) suggested the possibility that schizophrenics would show a greater tendency to inhibition than would non-schizophrenics. This inhibition was regarded as the explanation of certain of their characteristic catatonic symptoms. The inhibition was, in this case, considered to be protective, preventing the exhaustion of the nerve cells by bringing about a state of arrested activity. It was not a state of exhaustion or fatigue. The schizophrenic's defence of protective inhibition would presumably result in a loss of lability. This article reports an experiment which tests that inference. It cannot, however, be said with certainty that the experiment tests this aspect of Pavlovian theory because Pavlov held several slightly different theories about excitability, mobility and equilibrium. According to Teplov's (1958) recent review of Pavlov's view of types of nervous system, the conflicts between these do not seem to have been adequately resolved.

In a previous article the method of experimentation by verbal conditioning was described but a briefer description is also given below.

II. METHOD

Subjects. Three samples were compared: 20 normal males aged 24–44 years, $\bar{X} = 42{\cdot}4 \pm 12{\cdot}70$; 20 non-paranoid chronic schizophrenic males aged 27–55 years, $\bar{X} = 41{\cdot}25 \pm 6{\cdot}57$; and 20 chronic paranoid schizophrenics aged 36–55 years, $\bar{X} = 47{\cdot}5 \pm 4{\cdot}84$. The patients were, with few exceptions, those used in a previous conditioning experiment (O'Connor & Rawnsley, 1959) and had been selected on the basis of Bleuler's diagnostic criteria. Chronic was defined as two or more years' residence in hospital. The normal control group was composed of hospital staff randomly selected from a group of approximately the same age range as the patients. Each diagnostic group was subdivided into two matched groups on the basis of their conditioning scores in the previous experiment and allotted to a 'verbal' or 'visual' signal group.

Apparatus. The apparatus consisted of a circuit in which a control panel in one room enabled the experimenter to present direct vocal signals or light signals to the subject in the next room. The signals instantaneously operated a one-hundredth-second stop-watch and the subject's response, a hand pressure on a knob, stopped the watch and simultaneously recorded the pressure used in the motor act. The original conception of the experimental instrument has already been referred to (O'Connor & Rawnsley, 1959).

Procedure. All subjects not familiar with the apparatus were given conditioning trials, as in the previous experiment, several days before the lability trials were to be carried out. Thereafter, procedure was as follows. The subject sat at the display panel in the subject's room and was told to rest his hand on the desk in such a way that his fingers covered a telegraphist's key. He was told that everything would become clear and it was not necessary to give him any instructions. The experimenter then went into the other room saying that the subject would be able to hear him quite clearly and that the experimenter could see the subject through the window between the two rooms. An arrangement of screens, however, made it rather difficult for the subject to see the experimenter.

The experimenter then presented the subject with a signal, either a red light in one subgroup or the word 'light' in the other. After two seconds the experimenter said 'Press' if the subject had not already done so. When the subject had made the appropriate response with or without prompting the experimenter said 'Good'. The times of presentation of stimuli were randomized between twenty and forty seconds. Trials continued until the subject had made ten correct responses without prompting or until the subject had reached his fifteenth unsuccessful trial, and had been given a general instruction and subsequently made ten correct spontaneous responses. The general instruction took the form 'Whenever you see the red light (hear the word 'light') press the button'. This completed Stage 1.

Stage 2 began as a continuation of Stage 1, except that when the subject pressed in response to the experimenter's signal the experimenter uttered the following words, 'No, don't press'. This instruction was repeated as had been the word 'Press' in Stage 1, as often as necessary. The criterion was as before, ten 'successful' responses, that is ten responses in which the signal was not followed by a pressure of the subject's hand.

Stage 3 followed Stage 2 as this had followed Stage 1, without interruption. Failure to respond, correct in Stage 2,was now incorrect, and was followed two seconds after the signal, by the experimenter uttering the instruction 'Press' as in Stage 1. In both Stages 2 and 3 general instructions like those in Stage 1 were issued as necessary. The criterion in Stage 3 was the same as that in Stage 1.

The three stages might be referred to for convenience as C_1, conditioning, the first stage, Ex or extinction, the second stage, and C_2 the third stage or reconditioning. The use of such titles and their abbreviations may be question begging but they are used here simply for convenience of description. The first stage referred to as conditioning is, of course, reinforcement or reconditioning, and unlike Stage 3 does not follow an extinction stage.

III. RESULTS

Lability

The results are presented first as comparisons between diagnostic and signal groups in two-way analyses of variance. That is, diagnostic groups including the controls are compared with each other on their responses to both visual and auditory (verbal) stimuli. The types of measures used are difference measures because we are interested in 'lability', that is, the ease with which a positive stimulus becomes a negative one and vice versa. This is done by comparing the quantities (C_1–C_2), (C_1–Ex) and (Ex–C_2) as scores by diagnoses and signal type. This could be done with the number of trials needed to condition to the criterion, the latencies of successful trials in such scores as (C_1–C_2) and the pressures used in these trials.

Results of the conditioning measures of lability are presented first. These are made clear in Table 1.

Two-way analysis of variance showed no significant F ratios for diagnostic groups, but in two of the three cases, i.e. for the C_1–Ex and the Ex–C_2 measures, F for the difference between types of signal irrespective of diagnoses reached statistically acceptable levels of significance. They were in fact 20·20 ($P = 0{\cdot}001$) and 7·76 ($P = 0{\cdot}01$) respectively. C_1, Ex and C_2 scores, compared in a similar manner, naturally confirmed this difference between types of signal and also failed to reveal a difference between diagnostic groups. It should be noted that the difference was primarily due to the relatively low extinction score obtained by all groups in the non-verbal signal procedure, i.e. approximately three trials as compared with six for the verbal signal.

Table 1. *Measures of lability based on conditioning scores*

(I.e. number of trials to condition to criterion.)

Type of signal	Paranoid group			Non-paranoid group			Normal group		
	C_1	Ex	C_2	C_1	Ex	C_2	C_1	Ex	C_2
Verbal	5·4±4·9	6·4±4·6	7·5±6·8	4·3±4·6	6·0±4·6	7·2±5·5	2·5±4·6	6·8±4·6	9·2±4·9
Non-verbal	10·0±4·9	2·5±3·3	11·3±5·5	10·0±6·8	2·7±4·6	12·6±6·2	8·6±4·6	1·8±0·7	8·0±4·6

C_1 = conditioning; Ex = extinction; C_2 = reconditioning.

Table 2. *Measure of lability based on latency scores*

Type of signal	(C_1–C_2), paranoid group	(C_1–C_2), non-paranoid group	(C_1–C_2), normal group
Verbal	0·291±0·273	0·226±0·205	0·046±0·099
Non-verbal	0·067±0·331	−0·021±0·670	0·098±0·137

When the latency of response scores are considered in the same way only one score is possible, namely C_1–C_2, because during successful extinction trials (Ex) no latency scores are recorded. The latencies of unsuccessful trials might have been used but the number varied from subject to subject, making any selection procedure suspect. Table 2 presents the relevant latency scores. In some cases these scores are based on eight or nine subjects instead of ten because of apparatus difficulties.

Two-way analysis of variance showed that neither the differences between diagnostic groups, nor between types of signals affected response latency in this table. This does not mean that latencies of response are the same for the subgroups or types of signals. Individual tests on C_1 scores and C_2 showed differences between normals and both schizophrenic groups both for the visual and auditory (verbal) signals. These results are not presented because they do not concern us here and have been frequently demonstrated. The results presented, however, show that subgroups did not differ in lability as measured by the latencies of the (C_1–C_2) score although P approached the 0·05 level.

Pressure scores were also available for C_1 and C_2. These data are presented in Table 3.

Two-way analysis of variance of the score (C_1–C_2) showed no significant differences either by diagnosis or by type of signal or through their interaction. These results are therefore similar to those for number of trials to condition and latency of response in so far as they do not show any difference of lability between the diagnostic groups.

Variability

The variability of scores of latency and of pressure were expected to be greater for schizophrenics than for normals on the basis of the wide fluctuations of attention shown by many chronic schizophrenic patients. A test of this hypothesis was therefore made by comparing variances by subgroups and by type of signal on both latency and pressure scores. Results are given in Table 4.

The results show that variances are smaller for chronic schizophrenics than for normal controls in the case of pressure and larger in the case of latency. Two-way analysis of variance showed that pressures differed significantly between the groups. $F = 6{\cdot}37$, $P = 0{\cdot}01$ and latencies also differed significantly at a lower level of significance. $F = 4{\cdot}887$, $P = > 0{\cdot}05 < 0{\cdot}01$.

Table 3. *Measure of lability based on pressure scores*

Type of signal	(C–C), paranoid group	(C–C), non-paranoid group	(C–C), normal group
Verbal	0·133 ± 0·205	0·132 ± 0·195	0·116 ± 0·338
Non-verbal	0·030 ± 0·221	0·184 ± 0·644	0·094 ± 0·332

Table 4. *Mean variances for groups and type of signal*

Type of signal		Mean variance of C_1		
		Paranoid group	Non-paranoid group	Normal group
Verbal	Latency	0·108 ± 0·195	0·136 ± 0·151	0·026 ± 0·022
	Pressure	0·019 ± 0·028	0·036 ± 0·072	0·146 ± 0·142
Non-verbal	Latency	0·054 ± 0·079	0·190 ± 0·243	0·006 ± 0·006
	Pressure	0·026 ± 0·028	0·047 ± 0·054	0·115 ± 0·135

IV. DISCUSSION

The results would appear to leave no doubt that whereas Pavlovian conditioning theory would suggest a difference in lability between schizophrenics and normals, such a result was not found. Chronic schizophrenics are not less able than normals to convert a positive to a negative reflex and vice versa, i.e. under the circumstances prevailing during this experiment the performance of the schizophrenics was no less labile than that of the normal controls.

No statistically significant differences exist therefore between chronic paranoid schizophrenics, chronic non-paranoid schizophrenics, and normals so far as lability of conditioning and extinction are concerned: but examination of Table 1 shows that there is a clear difference between the ease of extinction according to the type of signal, whether visual or auditory (verbal). With a verbal signal, extinction is more difficult, i.e. more trials are needed to achieve it. There is no existing explanation of this rather striking phenomenon, but it may be related to another phenomenon discussed by Luria (1956), namely, the 'directive function' of language. Luria says: 'Motor responses, reflected and formed in the verbal system, are quickly formed, are no longer dependent upon continual reinforcement and do not disappear when there is a change in the manner of presenting the stimuli, do not break down in the presence of change or external inhibition, in short, they show all the signs of that stable and active self-regulating system which is created in close partnership with language.'

'Connexions which develop without the intimate association of speech are also extremely unstable. They depend upon constant reinforcement and disappear immediately reinforcement ceases'. This compelling quality of words might explain why change from a positive to a negative response might be more difficult with a verbal signal. The presentation of a non-verbal signal only facilitates one kind of reversal, i.e. from conditioning to extinction and not the other from extinction to reconditioning. Luria's explanation presumably applies only to the impelling and not the inhibiting function of language. It may be the case that easy inhibition is prevented by the use of words, whereas conditioning or reconditioning is relatively facilitated. The figures in Table 1 might justify such a point of view but we would need further experiments to support a conclusion of this kind.

Turning again to the question of lability or reversal capacity, it is necessary to note that such authors as Kutin (1956) and Astrup (1957) have drawn attention to the existence of pathological inertness and deficiencies of functioning of inhibition in residual or chronic schizophrenics. For the most part, they have not made predictions concerning lability and the results presented above would therefore appear to extend the work of these authors.

Our thanks are due to Dr E. P. H. Charlton, Physician Superintendent of Banstead Hospital, who kindly provided facilities for the experiments reported here.

REFERENCES

ASTRUP, C. (1957). Experimental research about disturbance of the higher nervous functions in defect schizophrenia. *Psychiat. Neurol. med. Psychol., Lpz.*, **9**, 9–14.

KUTIN, V. P. (1956). The problem of the pathophysiological characteristics of the chronic schizophrenic conditions. *J. Higher Nerv. Activity. (Pavlov)*, **6**, 251–9.

LURIA, A. R. (ed.) (1956). Problems of the higher nervous activity of normal and abnormal children. *Academy of Pedagogical Sciences R.S.F.S.R. Moscow*, **1**.

O'CONNOR, N. & RAWNSLEY, K. (1959). The acquisition of two types of conditioned response. *J. Abnorm. Soc. Psychol.* **58**, 157–61.

PAVLOV, I. P. (1927). *Conditioned Reflexes.* Trans. G. V. Anrep. London: Oxford University Press.

PAVLOV, I. P. (1938). *Twenty Years of Objective Study of the Higher Nervous Activity (Behaviour) of Animals*, 6th ed. Moscow: State Biological and Medical Publishing House.

PAVLOV, I. P. (1941). *Trial Excursion of a Physiologist in the Field of Psychiatry in Conditioned Reflexes and Psychiatry.* Trans. and ed. W. Horsley Gantt. New York: International Publishers.

TEPLOV, B. M. (1958). Study of the typological properties of the nervous system and their psychological expression. *S.C.R. Soviet Psychol. Inform. Bull.* **5**, 2–3.

(*Manuscript received* 20 *June* 1958)

THE BLOCKING EFFECT OF BENACTYZINE HYDROCHLORIDE ON A BEHAVIOUR DISTURBANCE INDUCED WITH LYSERGIC ACID DIETHYLAMIDE

By PETER H. GLOW

Medical Research Council Group for the Experimental Investigation of Behaviour, Department of Psychology, University College London

The main interest with lysergic acid diethylamide (LSD-25) lies with its capacity to mimic most of the clinical features of schizophrenia. It gives rise to short, sharp and transient psychotic disturbances in man and severe behaviour anomalies in animals. These behavioural disorders lend themselves to a systematic analysis. Some consideration has been paid to the development of blocking agents, that is, drugs which can reduce the severity or completely abolish the behavioural effect of LSD-25.

Experiments are described at three levels of inquiry at which it can be shown that benactyzine hydrochloride has a blocking effect on the behavioural disturbance in rats treated with lysergic acid. These experiments were conducted at the physiological, behavioural and pharmacological levels and the observations made are consistent.

I. INTRODUCTION

It was found in experimental work previously reported (Glow, 1957) that the LSD-25 disturbance in rats yields systematic relationships, not only between the various dosages and the intensity of their effects, but also between the effects and some physiological co-ordinates. If one assumes that LSD-25 has its effect first at the biochemical level, it is theoretically possible to vary systematically the biochemical factors involved, and thereby attempt to expose the underlying mechanism. One such approach is to search for compounds which have the property of reversing the disturbance, and then investigate these compounds and their action *per se*.

It is not unreasonable to commence one's search for compounds antagonistic to LSD-25 from that large body of drugs which have been found in recent years to be efficacious in controlling clinical disturbances. There are on record several attempts to use such ataractic or tranquillizing drugs for the specific purpose of reversing LSD-25 disturbances. None of these compounds has been found to be more than dubiously useful.

A compound which has, as yet, not been investigated with this purpose in mind is benactyzine hydrochloride. Benactyzine, the hydrochloride of benzilic acid diethylaminoethylester, was first synthesized in a systematic search for parasympatholytic compounds. The general pharmacology of the drug was investigated by Larsen (1955), who found the drug to have a low acute and chronic toxicity. Benactyzine has a pronounced anti-acetylcholine effect when measured on excised tissue preparations and blood pressure. The drug has no adrenalytic effect, nor ganglionic blocking action. Jacobsen & Sonne (1955) investigated the C.N.S. properties of benactyzine on stress-induced behaviour in rats. These authors found when using an escape from shock conditioning technique, that when the conditioned sound stimulus was presented the rats became tense and immobile. The administration of benactyzine 'abolished the tension in the situation and normalized the behaviour'. Similar studies were made on cats, using an air-blast in conjunction with controlled feeding (Jacobsen & Skaarup, 1955). These authors concluded that benactyzine was effective in 'normalizing' behaviour. The recent introduction of this drug as a tranquillizing agent into the armoury of the clinician has met with a mixed reception. A number of authors have noted with approval its usefulness in the treatment of anxiety cases (Hargreaves, Hamilton & Roberts, 1957). Coady & Jewesbury (1956) found benactyzine to have a striking suppressor effect on normal EEG rhythm. They also suggested that the drug produced 'thought blockage, impaired concentration, slowness and clumsiness in carrying out complicated procedures, and a general flatness of affect'.

Two things were considered when estimating the potential suitability of benactyzine as an LSD-25 blocking agent. First, there has been considerable speculation as to the psychological nature of the LSD-25 disturbance. Many workers have come to the not unreasonable position where they assumed that LSD-25 behaviour arises from disturbances in the autonomic nervous system. The implications are that the drug-induced anxiety state is a derivative or secondary reaction, to a condition of altered neurohumoral balance, in particular an antagonism with cerebral 5-hydroxytryptamine (5-HT). Secondly, a number of pharmacologists and biochemists have been concerned to develop tests on peripheral or excised tissues which could yield evidence about the basic activities which certain psychotherapeutic compounds hold in common. In this regard benactyzine has shown some properties which are potentially interesting. Berger, Hendley & Lynes (1956) demonstrated that benactyzine has an antagonistic effect on the spasmolytic activity of 5-HT on excised peripheral tissue.

Both the biochemical and the pharmacological evidence so far discussed could lead to an assumption that benactyzine has some effect at the 5-HT receptor level. The consequence of this effect on 5-HT activity would be an alteration in some aspects of C.N.S. neurochemistry in a manner which might conceivably block the LSD-25-induced disturbance. It was considered that if such a line of reasoning were justified, some autonomic indicator could be examined as a preliminary study to any behavioural investigation. LSD-25 produces a variety of measurable autonomic disturbances of central origin, and it is conceivable that the action of any blocking agent could be detected at this level. For this purpose the effect of LSD-25 on the heat regulatory activity of the rat was investigated. The hypothesis investigated in this first experiment was that benactyzine has antagonistic properties on the hypothermic effects of LSD-25 and that any LSD-25 blocking agent mediates its effects via the autonomic nervous system.

(i) *Experimental design*

It is known that LSD-25 has a hypothermic effect on rat rectal temperatures, and that this effect is species-specific to the rat, occurring with a dosage range of approximately 30 μg./kg. to 1 mg./kg. Temperatures are lowered by 1·5–2° C. for between 1½–2½ hr.

During preliminary trials it was found that dosages of the order of 20 mg./kg. benactyzine had an anti-hypothermic effect on LSD-25-treated animals. However, there also appeared a not unexpected variation on subsequent temperature between the dosage of LSD-25 administered, and the time at which premedication with benactyzine occurred. It was found that benactyzine has no effect on body temperature in the rat *per se*. The effects of benactyzine on rat body temperature concurred with those found by Larsen (1955).

With these facts in mind the following design was used. The subjects were twelve male albino rats obtained from the same population used in the behaviour experiment reported later in this paper. The animals were randomly divided into four equal groups, and caged separately at a room temperature regulated at 20° C. +3° C.

Experimental groups were pre-treated with 20 mg/kg. benactyzine hydrochloride* at varying time intervals. The control group received pyrogen-free water in equal volume. All groups received 500 μg./kg. LSD-25 as the hypothermic agent. Rectal temperatures were recorded at quarter-hour intervals for 1½ hr. after administration of LSD-25. The experimental and control groups were pre-treated at given intervals as follows.

Experimental group A, pre-treated 2 hr. with benactyzine;
Experimental group B, pre-treated 6 hr. with benactyzine;
Experimental group C, pre-treated 18 hr. with benactyzine;
Control group D, pre-treated 18 hr. with pyrogen-free water.

Temperatures were recorded with a small clinical thermometer to the nearest 0·1° C. Rectal temperatures preceding the administration of LSD-25 were recorded twice, with an hour time-lapse between them.

(ii) *Results*

The mean data have been set out in Table 1. It can be readily recognized that the effects of benactyzine are such as to give a measure of protection to the hypothermic effects of LSD-25 administration. The data were subjected to analyses of variance, the comparisons being made on the lowest temperature score derived from each group, the lowest temperature occurring at time 5. This work has been set out in Table 2.

Table 1. *Mean rat rectal temperature* (°C.) (3 *rats/group*)

500 μg./kg. LSD-25

Treatment order

	Group A	Group B	Group C	Group D
Pre-experiment temperature	37·7	37·3	37·5	37·3
	37·7	37·3	37·5	37·3
Time 1	37·7	37·1	37·4	37·2
2	37·5	36·9	37·0	36·5
3	37·1	36·8	36·7	36·4
4	37·1	36·8	36·7	36·3
5	37·1	36·9	36·8	36·7
6	37·3	37·6	36·9	36·7

Table 2. *Table of analyses of variance*

Source of variation	Sum of squares	Degrees of freedom	Variation estimate
Between samples	1·20	3	0·400
Within samples	0·27	8	0·0337
	1·47	11	

$F_{3,8}$ = 12·0 significant at the 1% level.

Orthogonal comparison

2 hr. and 6 hr. pre-treatment groups, $F_{1,8}$ = 3·18 not significant.
Control and 18 hr. pre-treatment groups, $F_{1,8}$ = 7·20 significant 5%.
Control and 18 hr. and 2 hr. and 6 hr. groups, $F_{1,8}$ = 25·59 highly significant.

The analysis of variance produces a significant result at the 1% level. An orthogonal comparison between the groups shows that while there is no significant difference between the pre-treatment groups 2 and 6 hr., there is a highly significant difference between these treatment groups and the control group. The data have been graphed in Fig. 1.

The disconcerting feature of these data is the wide variation in starting temperatures between the groups, particularly in the 2- and 6-hr. pre-treatments. It was considered possible that the initial variation in temperature could have affected the rate of temperature loss after the administration of LSD-25. In order to test this possibility new and larger populations were subjected to the same experimental treatment. The data have been set out in Table 3. It will be seen that these results are even more favourable to the hypothesis as the starting temperatures are closely bunched.

---

* All the benactyzine hydrochloride (Lucidil) was kindly supplied by Smith and Nephew Ltd, Herts, England.

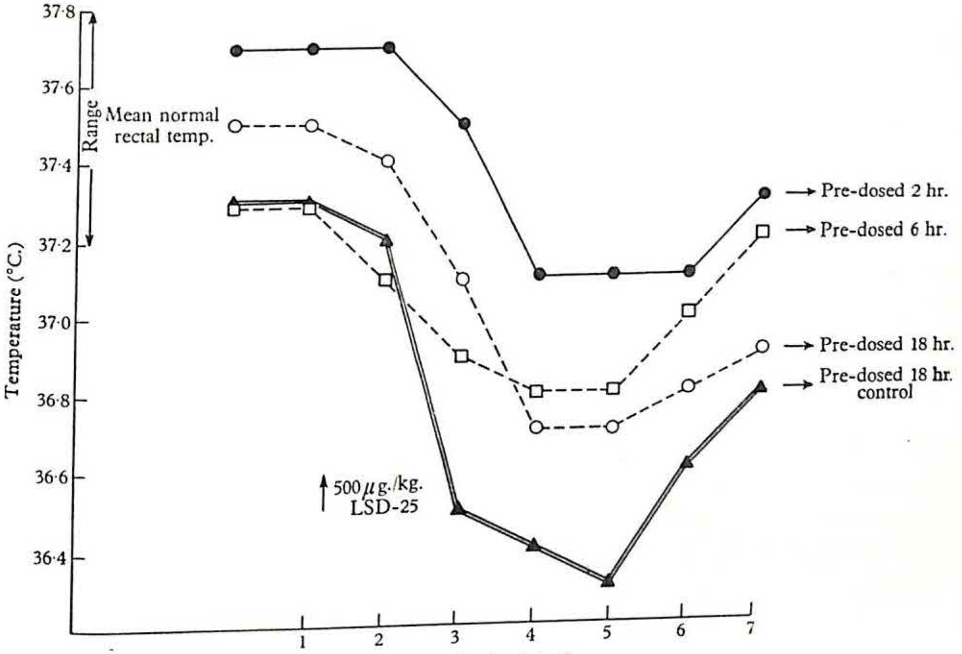


Fig. 1. Rectal temperatures after 500 μg./kg. LSD-25 (3 rats/group), effect of various pre-treatments with benactyzine hydrochloride.

Table 3. *Mean rectal temperature* (° C.)

Benactyzine, LSD-25

	Pre-treatment order		
	2 hr., 9 rats	6 hr., 9 rats	Control 6 hr., 6 rats
Pre-experimental temperature	37·5 37·4	37·5 37·7	37·6 37·7
Time 1	37·5	37·7	37·3
2	37·4	37·3	36·6
3	37·1	36·7	36·3
4	36·9	36·4	35·9
5	36·9	36·4	35·6
6	37·2	36·6	35·7

(iii) *Discussion*

The hypothesis as stated has been demonstrated to be reasonable. The evidence obtained tends to indicate that benactyzine alters the effect of LSD-25 on the central heat-regulating mechanism. It has also been possible to demonstrate that this effect has a relationship to the time of pre-treatment with benactyzine. The effectiveness of the pre-treatment decreases as the time elapses, before LSD-25 is administered. It is possible to show a log. to base three linear relationship between the pre-treatment with benactyzine time, and the mean body temperature recorded for the experimental groups. The lowest body temperature was recorded for all groups 1 hr. after the administration of LSD-25. The log.–linear relationship has been graphed in Fig. 2.

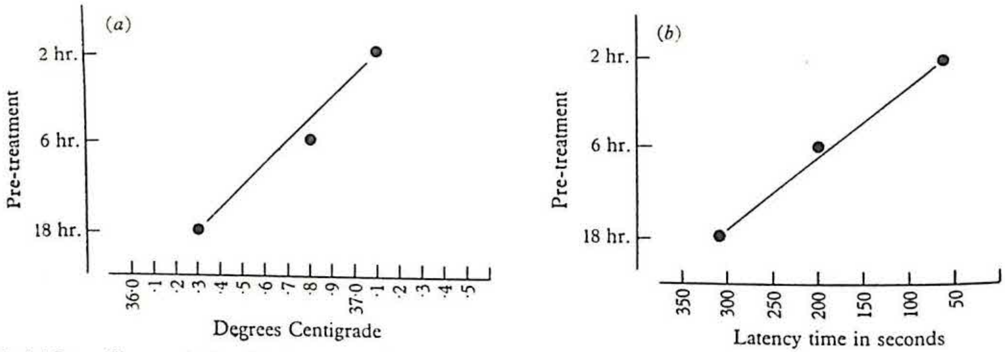


Fig. 2. (*a*) Log.–linear relationship between time of pre-treatment and mean lowest rectal temperature. 20 mg./kg. benactyzine hydrochloride/500 μg./kg. LSD-25. (*b*) Log.–linear relationship between time of pre-treatment and mean highest latency. 20 mg./kg. benactyzine hydrochloride/500 μg./kg. LSD-25.

II. Interaction between benactyzine and LSD-25 on behaviour

It has been possible to demonstrate that the effect of benactyzine on LSD-25-induced hypothermia is such as to tend to reduce the effects of this autonomic disturbance.

The relevance of this finding rests on the assumption that disturbance in autonomic activity is involved in the behavioural decrements associated with LSD-25 administration. Depending on the existence of some relationship between autonomic function and the observed behaviour, it is possible to surmise that benactyzine could be instrumental in reversing, or at least ameliorating the behavioural components of the LSD-25-induced disturbance. It can be speculated from the findings in the foregoing experiment that in this event benactyzine should be effective according to a given range of time/dose variability. The hypothesis for the following experiment was that benactyzine could be expected to ameliorate the intensity of the LSD-25 disturbance as determined by a response latency measure. Such effects as benactyzine may have should be related to the pre-treatment time at which the drug is administered.

Use was made of the long runway technique specially developed for the study of LSD-25 previously reported (Glow, 1957). The technique permits the measurement of response latencies and locomotor times.

(i) *Experimental design*

The experimental design was such as to permit a concomitant variation of the independent variable. The experiment was conducted with five replications. Each replication consisted of a litter of four male rats, one animal being randomly assigned to each experimental treatment. The equipment, handling, training and experimental procedure were identical with those used hitherto by the author. The subjects were obtained at weaning, and were of the same stock as all other animals so far reported. The pre-treatment times of benactyzine administration were as follows.

Experimental group I, 20 mg./kg., 18 hr.
Experimental group II, 20 mg./kg., 6 hr.
Experimental group III, 20 mg./kg., 2 hr.
Experimental group IV, water, 18 hr.

The dosage of benactyzine employed was determined on a series of pilot runs. It was found that at this dosage level no detectable alteration in either latency or running time occurred. The administration of a dose as relatively high as 20 mg./kg. was thought justifiable, as any results would then be comparable with those found in the previous experiments. A standard dose of 500 μg./kg. LSD-25 was administered to all rats. This dosage level can also permit a further comparison with the results recorded on the same dosage level in previous work with this drug. All injections were intraperitoneal. Latencies and running times were recorded at the following times: 5, 10, 20, 30, 45, 60 and 90 min. after the administration of LSD-25. In all cases the volume injected was made up to 2 c.c. A noise generator was used in order to induce running if latencies exceeded 5 min.

(ii) *Results*

The mean latency and running-time data have been set out in Table 4. The data has been graphed in Fig. 3 (latency) and in Fig. 4 (running times). It can be seen by inspection that there are highly significant differences in the latency scores relative to the temporal proximity with which the pre-treatment dose of benactyzine was administered. That is, the shorter the time interval between benactyzine administration and LSD-25, the less intense appears to be the disturbance as measured by latencies. There are, however, some difficulties in demonstrating levels of significance, if this is thought to be necessary. These difficulties arise out of the dependence that exists between successive observations, and the differential effect of the intense auditory stimulus administered, when latencies exceeded 300 sec., as occurred within the 18-hr. pre-treatment group.

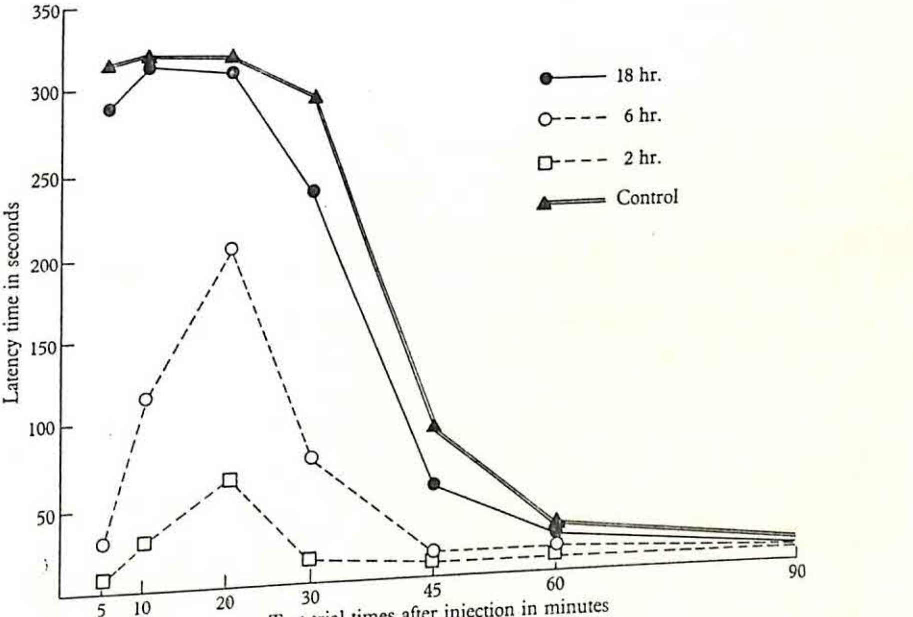


Fig. 3. Mean latency, all groups pre-treatment 20 mg./kg. benactyzine hydrochloride/500 μg./kg. LSD-25.

To circumvent these objections, a test of significance between the total latency times of the animals in experimental groups II and III was undertaken. These two groups did not require the use of the noise generator. Using a *t*-test for unrelated means $t = 123{\cdot}8$ with 8 D.F. significant at the 1% level. By inspection of the data and graph it can be seen that the systematic differences between treatment orders were directly attributable to the influence of benactyzine on the course of the LSD-25 disturbance. A log. to base three linear relationship can be found between the time of pre-treatment with benactyzine, and the mean latency decrements induced by LSD-25 (see Fig. 2).

The effect of LSD-25 on locomotor behaviour was not as marked as in the case of the latencies. The running times of the group subjected to the 18-hr. pre-treatment were faster than those of the groups pre-treated 2 or 6 hr. This effect is probably not due to the benactyzine, but rather to the intense auditory stimulus presented to this group.

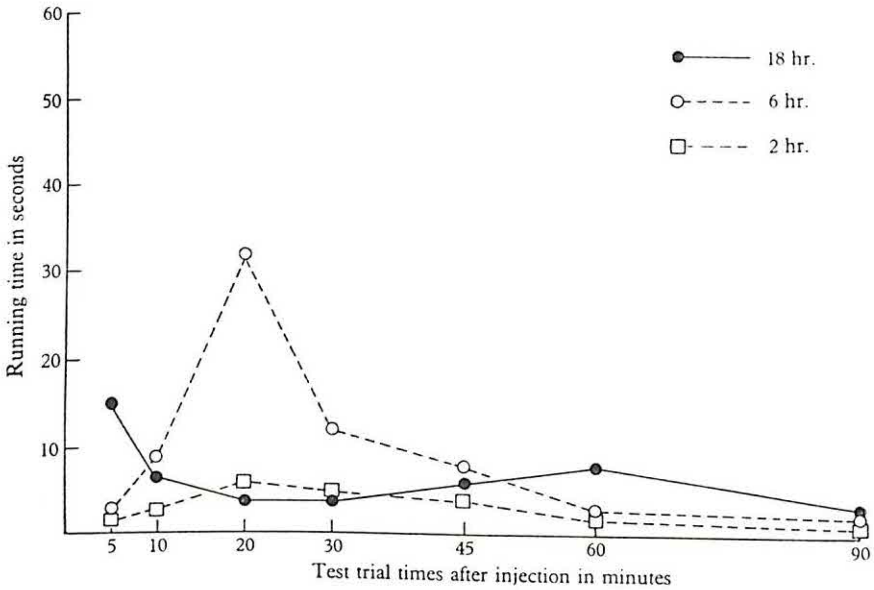


Fig. 4. Mean running time, all groups pre-treatment 20 mg./kg. benactyzine hydrochloride/500 μg./kg. LSD-25.

Table 4. *Mean latency and running time*

Benactyzine, LSD-25

	Pre-treatment order							
	18 hr.		6 hr.		2 hr.		Control	
Trial time in minutes	Latency	Running time	Latency	Running time	Latency	Running time	Latency	Running time
5	281	15	57	3	4	2	315	11
10	316	6	137	9	27	3	320	9
20	315	5	212	33	63	7	319	8
30	239	5	78	12	16	6	294	46
45	53	7	12	7	11	4	78	14
60	9	8	7	4	4	2	21	13
90	6	3	4	2	5	3	7	4

All figures taken to the nearest whole second.

(iii) *Discussion*

The onset rate of the LSD-25 action appears not to be grossly affected by any of the pre-treatment times. In all cases it would seem that the latencies are measurably lengthened. The degree of lengthening in the latency times is probably a function of the intensity of the total drug effect. It seems that if 'intensity' is a reflexion of LSD-25 activity in the C.N.S., as was thought could possibly be the case, then benactyzine has a central antagonistic effect on some LSD-25 actions.

If one compares the intensity effect found in this experiment with the control it would appear that 18 hr. pre-treatment does not alter the course of the LSD-25 disturbance. However, pre-treatment at 6 hr. reduces the intensity effect below the level found with 62·5μg./kg. LSD-25. Pre-treatment at 2 hr. almost completely ameliorates the LSD-25 disturbance. In all cases the height of the disturbance occurs 20 min. after the administration of LSD-25. It would appear that benactyzine pre-treatment does not alter the fundamental time relations in respect to onset, duration and recovery, as found in the author's previous report, but rather affects the intensity.

The observations made with LSD-25 on locomotor behaviour (Glow, 1957) appear to apply in this experiment. The effect of noise is such as to produce a locomotor response almost identical with that found in normal rats in this situation. In the case of the two pre-treatment orders where no criterion latencies were reached (i.e. 300 sec.) the noise generator was not employed. In the case of the pre-treatment 6-hr. group, running times are significantly increased as rats moved rather ponderously to the food box. Pre-treatment at 2 hr. shows a remarkable correspondence to normal untreated rats, excepting the trial at 20 min. after LSD-25 was administered.

The conclusions drawn from previous experimental work with LSD-25 were such that it could reasonably be assumed that the drug affects normal biochemical activity at specific, that is axodendritic, synapses. Taking this position as a plausible starting point, one may view the action of benactyzine as a specific LSD-25 antagonist. It is possible for instance to consider that benactyzine molecules readily take up positions at the tissue receptor level in a manner which excludes the LSD-25 molecule, but permits accessibility of the normal neurohumoral substance 5-HT, the excluded molecule being removed by enzymatic action.

This argument implies that the disturbance observed is related only indirectly to the biochemical action of LSD-25, and that the causal component is active 5-HT. This position would be in conformity with the observations made in experiments reported by the author (Glow, 1959). It was found that the duration–intensity of the LSD-25 disturbance was augmented under conditions where free 5-HT levels were raised. If one accepts this as a premiss it could follow that under conditions where the formation of active 5-HT is prevented, by the preferential exclusion of LSD-25, the incidence of an acute disturbance is reduced. One way of investigating whether such a model of 5-HT–LSD-25 activity is tenable would be to return from the molar behavioural level to the molecular level of the pharmacologist.

III. INTERACTION *IN VITRO* BETWEEN BENACTYZINE AND LSD-25

Unfortunately the biochemical and pharmacological aspects of benactyzine action have as yet not been adequately examined. Berger *et al.* (1956) demonstrated that benactyzine has an anti-spasmodic effect (in concentrations of 1·0/μg./ml.) on excised preparations of rat colon, when the spasmogen was 5-HT. This experiment was conducted on rat colon suspended in Ringer's solution at 37° C. These authors found that the spasmolytic action of 0·05–0·10μg./ml. 5-HT creatinine sulphate was completely blocked by 1·0μg./ml. benactyzine. Evidence has been accumulated that the rat uterus brought into oestrus is the most reliable and sensitive assay method for 5-HT and its action in the presence of antagonists (Page, 1954; Erspamer, Correali & Fiore-Donati, 1956). On these grounds it was considered necessary to conduct an investigation on a suitable preparation to detect the possible interaction pattern between the drugs under discussion.

The object of this assay was based on the notion that the blocking action of LSD-25 on the spasmolytic activity of 5-HT could be prevented by the administration of a suitable concentration of benactyzine hydrochloride.

(i) *Method*

Female rats (100 days old) were used 24 hr. after subcutaneous injection of 100 μg./kg. stilboestrol. All rats were killed by concussion and the uterine horns removed under de Jalon fluid. A longitudinal strip of uterus (2–3 cm. long) was attached at the bottom of a 5 ml. organ bath kept at 29 ± 1° C., containing de Jalon (1945) perfusion fluid (de Jalon fluid per litre: NaCl 9 g., KCl 0·42 g., $CaCl_2$ 0·06 g., $NaHCO_3$ 0·5 g. and glucose 0·5 g.). The other end of the tissue strip was attached to a balanced frontal lever, giving approximately five times magnification. Contractions were recorded on a smoked kymograph drum. It was found in practice that excised uterine tissue preserved in de Jalon for 24 hr. at 0° C. was more reliable than fresh tissue, giving fewer spontaneous contractions and maintaining its base line for several hours. The bath was continuously oxygenated.

(ii) *Procedure*

The following procedure was adopted during a 5 min. contraction cycle. The preparation was permitted contact for 1 min. with the spasmogen which was then washed out and replaced with fresh de Jalon fluid. The antagonist was permitted contact for 1 min., then washed out and replaced with fresh de Jalon fluid.

(iii) *Dose*

It was found that 25 μμg. 5-HT creatinine sulphate* in a 5 ml. bath produced submaximal contractions of a consistent height. The spasmogen was added to the bath in volumes of 0·2 ml. every 5 min.

(iv) *Results*

A dose of 50 μμg. LSD-25 was found to be maximally effective on blocking the third or fourth 5-HT-induced contraction, this being 15–20 min. after the administration of LSD-25. Thereafter recovery was slow but regular, until normal contractions were reached between 1–1½ hr. (see tracing in Fig. 5).

Benactyzine was added to the bath in concentrations varying between 100 μμg. and 2 μg. in volumes of 0·4 ml., and caused no contractions if given in the place of the spasmogen. Benactyzine had no spasmolytic properties in any dosage employed. If administered 1 min. prior to 5-HT (as in the case of LSD-25 above) it appeared to cause a slight diminution of the contractions, and occasionally an increase in spontaneous movements of the uterus. The effect of 2 μg. benactyzine is shown in the tracing in Fig. 5. It was found on a number of preparations that if the contact time with benactyzine (2 μg. in 5 ml.) was as long as 60 min., no further contractions could be induced with 5-HT. When 50 μμg. in 5 ml. LSD-25 was followed by benactyzine there was no noticeable effect on the subsequent antagonism to 5-HT contractions. Variable doses were added under this procedure ranging from 100 μμg. to 2 μg. in 5 ml.

When benactyzine preceded LSD-25 it was found that some measure of protection developed. The most effective concentrations for blocking were: benactyzine 2 μg. in 5 ml. (contact 1 min.—washed out) followed by 50 μμg. LSD-25 (contact 1 min.—washed out).

---

* Serotonin creatinine sulphate was kindly supplied by Dr Richards, Abbott Laboratories Ltd, Chicago, U.S.A.

With these concentrations it was possible to block completely the antagonistic action of LSD-25 on 5-HT-induced contractions. This effect may be viewed in the tracing in Fig. 5. A brief summary of these data may be found in Table 5.

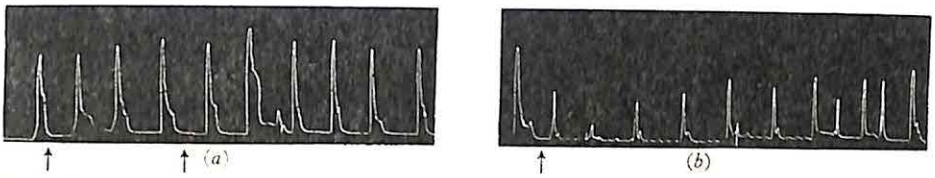


Fig. 5. Spasmolytic action of $25\,\mu\mu$g. 5-HT on isolated rat uterus. (*a*) $2\,\mu$g. benactyzine hydrochloride, followed by $2\,\mu$g. benactyzine hydrochloride + $50\,\mu\mu$g. LSD-25. Protection afforded to 5-HT contraction by administration of benactyzine hydrochloride with LSD-25. (*b*) $50\,\mu\mu$g. LSD-25. Inhibitory effect of LSD-25 on the spasmolytic action of 5-HT.

Table 5. *Interaction of* LSD-25 *and benactyzine on excised tissue*

Order of drug administration			
I	II	III	Results
1. 5-HT	—	—	Spasmolytic action
2. 5-HT	LSD-25	—	Blockade of spasmolytic action
3. Benactyzine	—	—	No spasmolytic action
4. Benactyzine	5-HT	—	5% decrease in 5-HT-induced spasmolytic action
5. 5-HT	Benactyzine	—	Spasmolytic action as in above 1
6. 5-HT	LSD-25	Benactyzine	Blockade of spasmolytic action as in above 2
7. 5-HT	Benactyzine	LSD-25	The spasmolytic action of 5-HT is protected from the effect of LSD-25

(v) *Discussion*

It was found that the antagonism of LSD-25 to 5-HT-induced contractions completely confirmed the observations made by Page (1954) and Erspamer *et al.* (1956). Benactyzine *per se* had no spasmolytic action on rat uterus, using a wide dosage range (1–20 units). When given before 5-HT-induced contractions it appeared to reduce the height at which the contractions usually occur, but no more than 5% at any time. It was quite usual, however, for the preparations to show increased spontaneous contractions. When benactyzine was left in contact with the tissue for an appreciable time it appeared to have antagonistic effects similar to those reported by Berger *et al.* (1956) on rat colon. It has also been found that with a contact time of 60 min. the preparations were no longer viable. The most interesting observation was that benactyzine, when given before LSD-25 in a dose ratio of $2\,\mu$g. in 5 ml./$50\,\mu\mu$g. in 5 ml., completely 'normalized' the 5-HT-induced contractions, i.e. obviated the LSD-25 effect.

IV. CONCLUSIONS

Benactyzine hydrochloride was used in order to determine whether this drug could be effective in reversing or ameliorating experimentally induced behaviour changes. The choice of the drug was made on the assumption that benactyzine had some action which reduces anxiety. LSD-25 has frequently been described as having an effect such that the resultant autonomic disturbances give rise to a state of anxiety. Whether in fact this is the manner in which the LSD-25 disturbance works was not fully investigated; rather only one aspect of its autonomic effects was isolated. It was found that 20 mg./kg. benactyzine gave a measurable protection to LSD-25-induced hypothermia. The degree to which this protection occurred depended on the time permitted to elapse between pre-treatment and the administration of LSD-25. Similarly, the latency times of rats in a runway were found to be normalized by the premedication with benactyzine. This observation could plausibly connect the autonomic and the behavioural effects of LSD-25.

Some consideration was given to the problem of the biochemical nature of the interaction between these two substances. A considerable amount of speculation and investigation has been concerned with the possible similarity of the peripheral antagonism of LSD-25 to 5-HT and the events occurring in the C.N.S. In view of the findings made with the benactyzine-normalizing effect on behaviour and body temperature, the relationships of these drugs were investigated at the *in vitro* level. It was found that benactyzine had 'normalizing' properties in this situation. A precaution must, however, be stated. It has always been held that the specificity of different tissues makes generalization difficult. This view must be especially emphasized when comparing results from excised peripheral tissue and the C.N.S. *in vivo*. The special circumstances existing in the C.N.S., carefully protected anatomically and biochemically, and in a specially discrete region at that (hypothalamus), make the circumstantial evidence provided by these experiments for such a C.N.S. process difficult to assess.

REFERENCES

BERGER, F. M., HENDLEY, C. D. & LYNES, T. E. (1956). Pharmacology of the psychotherapeutic drug benactyzine. *Proc. Soc. Exp. Biol., N.Y.* **92**, 563.

COADY, A. & JEWESBURY, E. C. O. (1956). Clinical trial of benactyzine hydrochloride as a physical relaxant. *Brit. Med. J.* **1**, 485.

ERSPAMER, V., CORREALI, P. & FIORE-DONATI, L. (1956). Pharmacodynamic and biochemical responses to 5-hydroxyhyptamine. *Arch. Int. Pharmacodyn.* **106**, 132.

GLOW, P. H. (1957). Some acute effects of D-lysergic acid diethylamide on the rat. *Quart. J. Exp. Psychol.* **9**, 206.

GLOW, P. H. (1959). Some aspects of the effects of acute reserpine treatment on behaviour, III. *J. Neurol. Neurosurg. Psychiat.* **22**, 24.

HARGREAVES, G. R., HAMILTON, M. & ROBERTS, J. M. (1957). Benactyzine as an aid in treatment of anxiety cases. Preliminary report. *Brit. Med. J.* **1**, 306.

JACOBSEN, E. & SKAARUP, Y. (1955). Experimental induction of conflict behaviour in cats: The effects of some anti-cholinergic compounds. *Acta Pharm. tox. Kbh.* **11**, 125.

JACOBSEN, E. & SONNE, E. (1955). The effect of benzilic acid diethylaminoethylester hydrochloride on stress-induced behaviour in the rat. *Acta Pharm. tox. Kbh.* **11**, 135.

LARSEN, V. (1955). The general pharmacology of benzilic acid diethylaminoethylester hydrochloride. *Acta Pharm. tox. Kbh.* **11**, 405.

PAGE, I. H. (1954). Serotonin (5-hydroxytryptamine). *Physiol. Rev.* **34**, 563.

(*Manuscript received* 10 *July* 1958)

ERRORS OF IMMEDIATE MEMORY

By R. CONRAD

Medical Research Council's Applied Psychology Research Unit, Cambridge

Problems of immediate memory have recently come under active consideration both by information theorists and by neurologists. The former have made use of quantitative facts, whilst the latter have based their views on some very general characteristics of forgetting. It is contended that the kind of changes that occur in short-term forgetting cannot be ignored in any account of immediate memory. Accordingly the main types of immediate memory errors are outlined, and some fresh material relating to known errors is introduced. A discussion is developed which considers the value of information theory and current neurological theories, in accounting for these phenomena.

I. Introduction

The main aim of this paper is to describe the nature of the commonly occurring errors of immediate memory, using digits as the source material. It is an odd observation that one of the greatest advances in our understanding of long-term memory processes came from the study of the character of imperfect recall (Bartlett, 1932), and not from the consideration of the quantity recalled which has dominated immediate memory studies. The experience of short-term forgetting is so common, the facts so easily demonstrated, and the time scale of the phenomenon so short, that the underlying processes must continue to present a challenge of high priority to the experimental psychologist. This paper then discusses the main errors that occur.

Formal interest in immediate memory for digits extends back to the work of Jacobs (1887). There followed a copious literature which Blankenship (1938) summarized. This literature is nevertheless concerned almost exclusively with two questions only. First, what conditions of presentation of stimuli determine the size of the span of immediate memory? Second, why are errors distributed so curiously? There is of course no limit to answering the first question, whilst Hull and others (1940) have proposed a solution to the second problem which has been widely accepted, but which cannot be directly tested. Perhaps these are sufficient reasons why, since immediate memory is always with us, the subject is as rich as ever it was.

In recent years, two trends have led to a renewed interest and fresh approach to old problems in this field. Concepts drawn from the theory of information have provided new ways of looking at the span and hence the underlying processes (Miller, 1951; Broadbent, 1957; Brown, 1958). Informational theories applied to memory are concerned mainly with the problems: how much information can be stored; why is information lost; which form does the loss take? In a general way much progress can be made along these lines, in the sense that a coherent picture can be drawn within the particular concepts and assumptions. In particular, the information theory approach has brought invigorating, if inconclusive, support for the theory of spontaneous decay of memory traces originally suggested by Müller. But the essentially quantitative character of information theory, its main attraction for psychologists, implies that experimenters tend to ask the questions that have always been asked, namely how much? The methods of analysing the derived quantities though have been different, and the conclusions drawn from them about mental processes have also been different. In general, from the time of Jacobs until the present day, emphasis has been on the proportion of all digit messages correctly recalled, and on the proportion of the digits in a single message which have been correctly recalled, though early work, much preoccupied with absolute span, gave weighted scores to different kinds of errors. In either case, errors were totalled and considered no further. The assumption made is that digits are either remembered or forgotten, whether subsequent analysis is based on information theory or on older concepts.

Pollack (1952), in an attempt to get at information which was only partially forgotten, informed his subjects when they were wrong in an item and invited further attempts at recall. Unfortunately Pollack does not indicate whether he treated transposition of two adjacent digits as one or two errors for this purpose. If the latter, subjects were inadvertently being misled. If the former, then they were being helped by the experimenter to defeat the scoring method.

The informational theory of memory has been considerably extended by Brown (1958) who has proposed that forgetting is a function of redundancy. An item can be said to carry redundant information when some of the information can be lost without the item losing its unique identity. As Brown has put it, a number chalked on a blackboard can be smudged to a certain extent before it becomes unrecognizable as a number. According to the theory then, the more redundant information in an item, the longer it will be remembered.

The second modern trend which has renewed interest in immediate memory problems derives from recent neurological theories of memory processes. Apart from the question of the anatomy of memory processes, Eccles (1953) in particular has reported details of plastic changes at synapses, earlier postulated by Adrian (1947), which have been considered to provide a link between a reverbatory circuit concept of immediate memory, and a permanent neuronic change concept of long-term memory. Whilst current neurological theory of memory is plausible, it is formulated in a manner which makes it impossible to test psychologically. But because this is possibly the fault of psychologists in not providing relevant data, a secondary aim of this paper is to try to fill in some gaps, perhaps thereby providing some clues for neurologists. In other words, a knowledge of some of the characteristic changes that occur during short-term memory storage might be of more value to neurological theorists than a knowledge of the quantity of change that takes place.

One of the difficulties for the psychologist is that whereas it is easy to define an input into memory storage in descriptive terms which are familiar, such as digits, no one thinks that digits as such are stored in the brain. The concept digit embraces a complex pattern of sensory stimulation which can be widely generalized into different sensory modes, different temporal and spatial patterns, and yet still be represented in memory as a unique item. For instance, the symbols '3' and 'III' and 'three' all mean the same thing, and a sound can be heard which also registers as the digit 3, and this sound can be a male or female voice sound, or can be uttered artificially by a machine. If a person is asked to repeat the digit, he does not reproduce exactly the original stimulus, but his own generalized version of it. Within certain limits he has no difficulty in preserving its conceptual identity. The critical neurological problem is: how is this identity represented in the brain, how do we select it on demand from the representations of all the other unique identities in the brain, how is it preserved, and why sometimes can it not be reproduced on demand?

Information theory skirts this difficulty by referring not directly to concepts or sensory patterns, but to information quantified in bits. The number of bits of information is determined for any particular item by the number of items in the class from which the particular item is drawn. The assumption is made that the class itself is clearly defined. For example, the digit 3 is one of ten digits, and the number 33 is one of a set of one hundred. The main assumption made in predicting performance from information theory, is that subjects are playing the game according to the experimenter's rules, i.e. the subject accepts, and is familiar with, the class of items from which the test material is drawn. From this, Miller (1951) has argued that the larger the class of items is numerically, the less likely is the subject to be able to store and reproduce (memorize and later recall) items from the class.

At the present time no systematic approach to problems of immediate memory, other than the two mentioned, seems likely to command serious attention. On the other hand, neither at the moment can offer anything like an adequate explanation of the main facts of memory behaviour. The following description of commonly occurring errors may provide a challenge to all theorists.

The material for the subsequent discussion comes from a number of experiments, all of which were designed to examine the effect on recall of different presentation and recall conditions. The details of these experiments need not concern us here, nor need the general trends of their results. Where the particular experimental conditions seem relevant to the kind of error which occurred, the fact will be noted and discussed. The general form of all the experiments was the same. Digit messages of varying lengths were recorded on magnetic tape. Subjects listened to the tape and after each message either wrote down what they remembered of it, or responded by dialling the number on a telephone dial or keying it on a two by five digital keyboard. Subjects were either Post Office telephonists, or Naval ratings, and there were never less than twenty-four in each experiment.

II. TRANSPOSITION ERRORS

The simple transposition of two adjacent digits referred to by Smith (1895) is the least ambiguous to interpret of all errors. It occurs frequently in messages of any length, and the probability of correct guessing is relatively small. Less often three adjacent digits in a sequence of perhaps eight are transposed, and occasionally four are transposed. By this point analysis becomes somewhat clinical. For instance consider the following case:

Input: 1 2 3 4
Output: 2 3 4 1 (*a*)
Output: 2 4 1 3 (*b*)

In case (*a*) the correct digits are reproduced, only one is sequentially misplaced, four digits are 'wrong'. In case (*b*), all four digits are misplaced. Calculation of probabilities to mean anything relevant to memory begins to involve assumptions which are far from clear. Fortunately most transposition errors are of two or three digits only; the transposition error is the only error in the message, and hence objective scoring is easier. Table 1 gives the frequency with which different numbers of digits were found to be involved in transposition errors with 8-digit messages.

In a series of 8-digit messages, clearly detectable transposition errors accounted for some 50 % of all wrong digits. It is of interest to note first, that subjects who make few errors in general, have a higher proportion of transposition errors than subjects who make many errors (see Table 2). This suggests at first sight that with the fading of a memory trace, the order elements are less stable than the item elements.

A second point of note about transposition errors, is that their distribution serially is exactly the same as the distribution of all other errors, namely U-shaped skewed towards the end. Like other errors, transpositions rarely occur right at the beginning or the end of a message. This suggests that the chance of an error at any particular serial position is independent of the kind of error likely to occur. It might be stated here that no evidence has been found that a particular kind of error is more likely to occur than any other kind at any particular position in the series.

Table 1. *Frequency of transposition errors*

2 digits transposed 58 times per 100 transposition errors
3 digits transposed 36 times per 100 transposition errors
>3 digits transposed 6 times per 100 transposition errors

Table 2. *Distribution of errors amongst subjects*

	Type of error	
	Transposition	Other
4 best subjects	36	31
4 worst subjects	203	340

$x^2 = 6{\cdot}68; \quad P > 0{\cdot}01.$

Information theory being essentially quantitative, it must, to be useful, predict quantities and quantitative change. But this prediction depends on being able to define an input and an output informationally. So far this has been done only with respect to 'number' in the case of immediate memory for digits (Miller, 1956; Pollack, 1952). But Brown (1958) has suggested that order information should be separately considered. This seems to be an important suggestion, but involves methodological difficulties of major proportions. If one is to treat order information loss in the same way quantitatively as item (number) information loss, then some kind of correlational analysis between input and output must be carried out for each individual message, in order to determine how much order information had been lost. Furthermore, if any item information is lost then the assessment of order information cannot be made at all because the only data on which it can be based will have gone.

Transposition errors are so common, striking, and frequently unambiguous, that they cannot be ignored. Nor has a neurophysiological mechanism yet been suggested which could possibly account for this phenomenon. Eccles (1953) points out that '...in continuously circulating round any particular chain of neurones, the impulses would give a specific pattern of activity in space and time, which would of course form the basis of specific sequalae, e.g. conditioned reflexes or memories'. But not only does this preclude the possibility that the temporal pattern could become disordered, but by linking memory with conditioned reflexes, positively suggests that it would not, since the development of a conditioned response depends specifically on rigid sequence of events.

III. Omission errors

Logically, an omission error represents a condition when the subject admits that he has forgotten. In testing, this would normally mean a blank space left in the record. But of course, the subject might forget and guess. Whether a physical gap is left partly depends on the kind of instruction given. If special attention is drawn to the possibility of omission, subjects may be less likely to guess when they have in fact forgotten. But the frequency of recorded omission in any case will depend on the method of reproducing the recalled message. If the message length is known to the subject, and he writes down his recalled version, or he reproduces it in a form that he can see, then recorded omissions are most likely to represent what has mentally occurred. But if, for example, *O* repeats back the message verbally, in order for him to locate omissions correctly, he must remember not only the message, but also the number of digits reproduced as he proceeds. If the message length is randomly varied as well, it will be impossible for the experimenter to know, in many cases, whether or not an omission was intended, and if it was, where in the series it was intended to locate it. Since, from the point of view of assessing recall, a guess is equivalent to an omission, the question can be considered further in the next section.

When an error of omission can be definitely established, one has a choice of several equally plausible theoretical explanations. Psychologically the memory trace has decayed in the absence of rehearsal. Informationally, noise has reduced the redundancy to the point where the digit retains insufficient information for identification. Neurologically, the excitation initiated by the stimulus has, in the absence of reinforcement, died away. When the subject admits that he cannot remember—that he has no idea of what the missing digit is—then we can safely classify the error as an omission. This is the first and most obvious phenomenon with which any theory of memory must deal. But if the subject reproduces any digit, even though it appears to be a wild guess, it cannot be recorded as an omission without careful analysis.

IV. Substitution errors

When a subject records a digit different from the digit which is correct for that position in the series, and it is clearly not part of a transposition, the mere statement that it is wrong may be inadequate for an understanding of memory. The status of a good many wrong digits will be ambiguous. This ambiguity merits discussion. If message lengths are such that subjects will have some error in about two or three messages out of ten, it is found that very many messages are wrong because there is a transposition, or because a single digit is incorrect. An examination of this latter easily defined category provides the data for what follows in this section.

This kind of error could fall into one of three classes. (*a*) A class of random error. The subject fully believes that his recall is correct, yet there is in fact no apparent relationship between input and output. (*b*) The subject has completely forgotten, is aware of this, and guesses. In effect this is an error of omission. These two classes are indistinguishable objectively. (*c*) Any systematic replacement of one digit by another. This would include 'acoustic confusion' between input and output, the possibility of which was mentioned by Meumann (1905) and given partial credit in measuring memory span. The idea of acoustic confusion is attractive with much face validity and a good deal of everyday experience to support it. Woodworth (1938), for instance, recounts how in trying to recall a name, the first incorrect attempt often bore an 'acoustic' relation to the original. But other systematic confusions are possible. For example, the subject might tend to replace forgotten odd numbers by other odd numbers. Clearly in any case, two kinds of confusion are possible, perceptual confusion and recall confusion.

Perceptual errors, especially those leading to acoustic confusions, are obviously likely when the listening conditions are unsatisfactory. Recall errors equally have been demonstrated by Bartlett (1932) for long-term recall, and Lashley (1950) has suggested a neurological basis for immediate memory which could account for systematic confusion between digits in immediate recall. In this case we would expect that two digits that had many neurones in common as the basis of their temporary storage, would be more likely to be confused. But we do not know which features of digits to consider relevant. Will like-vowelled digits such as 5 and 9 in storage share much common neuronic network? Or will initial consonant be critical as with 6 and 7? Or might contiguity in the ordinal scale such as 8 and 9 provide the basis for sharing neurones? There is no certainty about any of this.

Another theoretical basis for postulating the existence of recall rather than perceptual confusions is provided by informational theories of immediate recall. In this case one would say that as a result of decay, information is lost, and that what remains suggests another digit. Or that two digits which are distinct when all the information regarding them is present, are indistinct when either or both lose some information. This particular theory has the advantage over a neurological explanation, that, e.g., the digit 8 with some information lost might suggest 0, but that 0 with some information lost need not necessarily suggest 8. But as with Lashley's idea, the theory cannot really predict what will confuse with what, since we again do not know which physical features of the digit carry which information, nor do we know whether all the information is lost at a constant rate, e.g. 8 and 0 might confuse after a little loss, but not after a big loss. Nevertheless informational theory accounts of memory would be critically supported if evidence for systematic confusion of this kind could be found.

Evidence regarding the specific nature of digit confusions is largely anecdotal. Burt (1909) used Meumann's method of crediting confusions as partially successful recall, but he used nonsense syllables as material, and did not systematically analyse the nature of the confusion. More recent studies using digits have been concerned with the intelligibility of physical communication channels, and these have expressly introduced the likelihood of perceptual confusions.

It was against this background that data regarding digit confusions in recall were sought (Conrad, 1958). These data were taken from a series of 8-digit messages so constructed that each digit (0–9) occurred equally often in each serial position. This equated the probability of error of each digit, for it had been found that the digit which occurred most frequently within a series of messages, was most frequently substituted in error. Twenty-four telephonists took part in the experiment, but not all of the error data were used. The data selected for analysis were those where only a single digit in the message was wrong. This selection may have introduced bias into the results. But there is no other way of being quite certain that transposition errors are excluded. In all, 687 wrong digits were considered. In each case the correct and the substituted digit were known. The specific questions asked were: is any digit substituted for any other one digit more often than would be expected by chance? If so, were these reciprocal confusions, i.e. 4 for 8 *and* 8 for 4?

A confusion matrix for all ten digits was set up showing the frequency with which each digit was substituted for each other digit. An over-all χ-squared test showed that at the 5% significance level there were differences between observed and expected values. These differences were clarified by examining the distribution of the substitutions for each digit separately. This analysis revealed that there was only one case where one digit was substituted for another more often than chance expectation at the 5% level (at the 2% level, no discrepancy was different from chance). Digit 3 was too frequently substituted for 2, but the reverse substitution was not different from chance expectation. It is difficult to think of any unique principle which might underly this and only this substitution, so that the question whether substitutions are due to perceptual or recall error has little significance. In general there appear to be no systematic substitutions in a noise-free system, but merely chance arrangements. One must conclude that evidence for partial forgetting or partial item informational loss, is lacking. However, a casual observation on the data suggests that some substitutions may have a completely unexpected origin, which undermines the theory of random substitution, but without supporting that of partial forgetting.

V. SERIAL ORDER INTRUSIONS

Melton & Von Lackum (1941) were the first to report a further type of error of immediate memory. They used sequences of short lists of words, and found that recall errors were frequently words taken from the previous list *and* from the correct serial position. Examination of the present material suggested that many substitutions were in fact serial order intrusions. A message, for example, which ended ...258 was reproduced as ending ...948. The previous message correctly recalled, ended ...940.

To follow up this lead, some simpler data were analysed which were drawn from an experiment in which the subject heard messages of 3–6 digits which were immediately followed by an interpolated task of 4–12 sec. Recall of the original message was required at the end of the interpolated task. This experiment was very similar in form to that reported by Brown (1958). Even with the 3-digit message and 4 sec. of interpolation, the subjects, who were female telephonists, failed to recall correctly about 20% of the messages, and with 8 sec. of interpolation about 50%. It was these failures taken together which provided the data for the following analysis.

Inspection immediately revealed the intrusive character of errors, e.g. two successive messages 617, 483, might be reproduced as 617, 413. A 3 × 4 substitution matrix was set up permitting each wrong digit to be placed in a cell according to whether the digit which was substituted for it was the digit in the same serial position in the previous message, was the digit in either of the other two positions in the previous message, or was any other digit. The matrix is shown as Table 3. Table 3 is interpreted as follows. If the first digit of message 2 is wrong and is substituted by the first digit of message 1, then a record is made in cell *Aa*. If it is substituted by the second digit of message 1, it is recorded in cell *Ab*. If the substitution is a digit that does not occur in the previous message, it is recorded in cell *Ao*. Similarly for the other wrong digits in message 2.

This analysis was simplified by ignoring transposition errors and when this was done it was found that no digit was wrongly repeated in the same message. This was probably because subjects very quickly learned that the messages were so constructed that no digit did occur twice in a message. On the hypothesis, therefore, that a wrong digit would be substituted by any other digit, any wrong digit would occur with a probability 1/9. Thus in Table 3 the expected probability in each cell of rows *a*, *b*, and *c* would be 1/9 and in each cell of row *o*, 6/9 of the column totals.

Table 3. *Substitution matrix*

		Input		
		A	*B*	*C*
Output	*a*	26	2	3
	b	3	42	2
	c	1	2	23
	o (other)	21	17	14

Simple inspection shows that the actual distribution is very different. There are far too many observations in cells *Aa*, *Bb*, *Cc*. It is equally clear that a wrong digit is not substituted by any of the digits in the preceding message, but only by a particular one. Thus Melton and Von Lackum's result is fully confirmed.

Having established beyond doubt the occurrence of these errors, the 8-digit material was re-examined. What follows is necessarily a clinical interpretation, but given the results of the analysis of 3-digit messages, and the known facts of the character of errors of immediate memory described earlier, the classes of error found will be recognized as having a high face plausibility. The method was merely to examine messages which had some wrong digits, and to see whether a relationship could be reasonably observed with the digits recalled in the previous message. There was this general point. Errors were found to be related not necessarily to the previous message, but to the subject's reproduction of it. The following relationships have emerged:

1. Simple substitution of a single digit occurring in the same serial position in the previous message. This was illustrated earlier.
2. Substitution by erroneous digits from the previous response.

	Input	Output
Message 1	81394062	81043962
Message 2	68538429	68048429
Message 1	92468536	92468736
Message 2	09358672	09358732

It is particularly noticeable that this and other classes of error are more likely to occur if the arrangement of digits is such that by substituting one or two digits, two or three then repeat those in the previous message.

3. Substitution facilitated by transposition.

	Input	Output
Message 1	685....	685....
Message 2	608....	680....
Message 1	...3806	...0836
Message 2	...3586	...5836

4. Substitution facilitated by pattern.

	Input	Output
Message 1	0416	0416
Message 2	3856	3456

The four classes exemplified above cover the main kinds of serial order intrusion errors that have so far been found. There are also many cases where more than one class appears to be involved, and other data might yield further classes. The most immediate conclusion of significance from this analysis is that substitutions are not necessarily chance guesses as suggested by the confusion matrix discussed earlier. Substituted digits are distributed approximately equally amongst the possibilities partly because the messages were made up from tables of random numbers. In substituting by digits from the previous message, subjects were in fact substituting from random numbers. It should, though, be stressed that not all errors can be accounted for by the above intrusion error classes. The other kinds of error referred to earlier in this paper all occur independently of intrusion effects. It is not possible to say, at this stage, what proportion of all errors other than omissions are errors of serial order intrusion.

This analysis then poses the question as to whether intrusions of this kind actually create or predispose to error, or whether the occurrence of error is independent of intrusions; the intrusion nevertheless determining the nature of the error. The available data themselves cannot clarify this. Melton and Von Lackum felt that the intrusions they found arose from some kind of inhibition. In the context of verbal paired-associate learning, different from the Melton and Von Lackum, and from the present situation, Morgan & Underwood (1950) again associated inter-list intrusions with inhibition and suggested that the intrusion competed with the correct response. In both cases the link between intrusion and inhibition was inferred rather than demonstrated. For the present, therefore, the question whether serial order intrusions are a special case of proactive inhibition must remain unanswered.

It is, though, clear from the data on intrusion errors, that under the test conditions provided messages are surrounded by a good deal of disturbance which is very far from random, and which is generated by the previous response. This particular characteristic of the system is difficult to explain informationally. On the other hand neurological theory seems better able to handle this phenomenon. A facilitated pathway, facilitated through recent use, could presumably discharge when the correct pathway was deprived of stimulus. But just the same the dominant effect remains as much a neurological as a psychological mystery, and at this stage when relatively little is known about it, it is perhaps wisest to leave the problem open. But that it is a demonstrable fact cannot be evaded, and suggests that the encoding of groups remains flexible and relatively loose.

VI. CONCLUSION

Too much should not be read into the apparent comparison of the ability of informational and neurological theories to account for the errors of immediate memory which have been described. Information theory deals with abstract quantities which can be used with advantage to replace the loose idiom of everyday language about memory. But there is a limit to this re-definition, which may depend on no more than our present state of knowledge about sensory process. We do after all now have workable scales by which we can quantify degrees of illumination or sound intensity, when the layman is still obliged to refer vaguely to brighter and louder. We have no such scale for degrees of forgetting. This means that there is much uncertainty about what, and how much, is lost when there is a discrepancy between what is perceived and what is recalled. Some of the characteristics of this discrepancy have been discussed in this paper, and it has been argued that because we are not sure in which kind of way the changes might be quantified, so it would appear that the data cannot easily be assimilated into an informational account of forgetting.

The matter of neurology, on the other hand, is specific and physical. Neurological theory wields observed or postulated material processes, which must be correlated with the psychological processes of forgetting and remembering. When a correlation can be established between a psychological and a neurological change, the relative quantities of change are probably of less interest than the existence of the phenomenal relationship as such. But here a vicious circle situation predominates. The neurophysiologist complains that psychology does not provide him with enough incontrovertible data on memory behaviour. The psychologist complains that neurological theories of memory function are too vague to provide the basis for psychological experiments which might test them. If there is positive evidence (Burns, 1958) for synaptic facilitation resulting from repeated stimulation becoming some kind of structural and hence more permanent change, this could explain the simplest decay properties of memory traces which are everyone's common experience. No one, though, has attempted to explain neurologically how patterns of neurone excitation can lose their sequential character and acquire another, nor what property of a group of neurones can create the characteristic serial position effects of errors of immediate memory.

This is not merely a question of filling in the gaps. Once we have agreed to the general experience of forgetting, these are the main features of the process, and no theory can be considered adequate unless it can account for them.

REFERENCES

ADRIAN, E. D. (1947). *The Physical Background of Perception.* Oxford: The Clarendon Press.

BARTLETT, F. C. (1932). *Remembering.* Cambridge: The University Press.

BLANKENSHIP, A. B. (1938). Memory span: a review of the literature. *Psychol. Bull.* **35**, 1–25.

BROADBENT, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychol. Rev.* **64**, 205–15.

BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. Exp. Psychol.* **10**, 12–21.

BURNS, B. D. (1958). *The Mammalian Cerebral Cortex.* London: Edward Arnold.

BURT, C. (1909). Experimental tests of general intelligence. *Brit. J. Psychol.* **3**, 94–177.

CONRAD, R. (1958). Accuracy of recall using keyset and telephone dial and the effect of a prefix digit. *J. Appl. Psychol.* **42**, 285–8.

ECCLES, J. C. (1953). *The Neurophysiological Basis of Mind.* Oxford: The Clarendon Press.

HULL, C. L. *et al.* (1940). *Mathematico-deductive Theory of Rote Learning: A Study in Scientific Methodology.* New Haven: Yale Univ. Press.

JACOBS, J. (1887). Experiments in 'prehension'. *Mind*, **12**, 75–9.

LASHLEY, K. S. (1950). In search of the engram. In Soc. for Exp. Biology, *Physiological Mechanisms in Animal Behavior.* New York: Academic Press.

MELTON, A. W. & VON LACKUM, W. J. (1941). Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.* **54**, 157–73.

MEUMANN, E. (1905). Intelligenzprüfungen an Kindern der Volksschule. *Z. Exp. Päd.* **1**, 35–101.

MILLER, G. A. (1951). *Language and Communication.* New York: McGraw-Hill.

MILLER, G. A. (1956). The magical number seven, plus or minus two: some limits of our capacity for processing information. *Psychol. Rev.* **63**, 81–97.

MORGAN, R. L. & UNDERWOOD, B. J. (1950). Proactive inhibition as a function of response similarity. *J. Exp. Psychol.* **40**, 592–603.

POLLACK, I. (1952). The assimilation of sequentially encoded information. 2. Effect of rate of information presentation. Human Resources Research Laboratories, Memo Report No. 25, Air Research and Development Command, U.S.A.A.F.

SMITH, W. G. (1895). The relation of attention to memory. *Mind*, **4**, 47–73.

WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Henry Holt.

(*Manuscript received* 28 *July* 1958)

SERIAL POSITION EFFECTS IN NONSENSE SYLLABLE LEARNING AS A FUNCTION OF INTERLIST REST PAUSES*

By H. J. EYSENCK

Institute of Psychiatry (Maudsley Hospital), University of London

An experiment to test a consolidation hypothesis of certain serial learning position effects has been reported, in which effects produced by continuous and discontinuous presentation of lists were compared. The failure of these two conditions to produce different position effects disproves the hypothesis, but also argues against most of the currently held theories.

The experiment to be reported arose from a theory designed to explain one of the well-known serial position effects in nonsense syllable learning, namely the fact that the final syllables in a list are learnt more easily than intermediate ones, and almost as easily as the first few syllables in the list. The hypothesis derives from the *perseveration* or *consolidation* theory of memory, according to which memory traces require some time for their consolidation (McGeoch & Irion, 1952, p. 118). The time provided by the interlist rest pause would aid in the consolidation of the last one or two syllable memory traces, and might thus account for the superior memory for these syllables. A necessary deduction from this hypothesis would be that the abolition of the interlist rest pause would abolish the superiority of the last few syllables in the list. A previous study by Mitchell (1933) did not show the predicted effect, but as there are some doubts about the proper control of relevant variables, a repetition of the experiment seemed to be in order.

Two fifteen-syllable lists (A, B) were constructed and shown to the subjects on the Maudsley memory drum (Eysenck, 1957), which enables discrete 2-sec. presentations of each syllable to be made. In addition, a set of fourteen dual-colour rectangles was affixed to the proper positions on the memory drum terminating with a single-colour blue rectangle. Two types of presentation were used: (1) Continuous presentation, in which the first syllable followed the fifteenth without a break, and (2) discontinuous presentation in which each series was followed by colour naming, the last single colour serving as a cue for the first syllable. Eight subjects were employed, all of them post-graduate psychologists ignorant of the purpose of the experiment. Order of series, and continuity or discontinuity of each series, were counterbalanced for the subjects so that effect of series, subjects, and order were experimentally controlled. The results of the experiment are shown in Figs. 1 and 2, in which are plotted respectively the *mean first correct response* and the *total number of errors* made before one complete correct presentation of the list.

The results disprove the hypothesis conclusively, terminal position effects being as marked for the continuous as for the discontinuous presentation. The only difference between the two presentations appears to be an improvement in learning for the whole series under the discontinuous presentation condition; this is significant beyond the 0·01 level. The result, therefore, is similar to that of Mitchell. It appears to disprove not

---

* I am indebted to the Bethlem Royal and the Maudsley Hospital Research Fund for support of this investigation. I am also indebted to Miss S. Kurlender for help in the investigation.

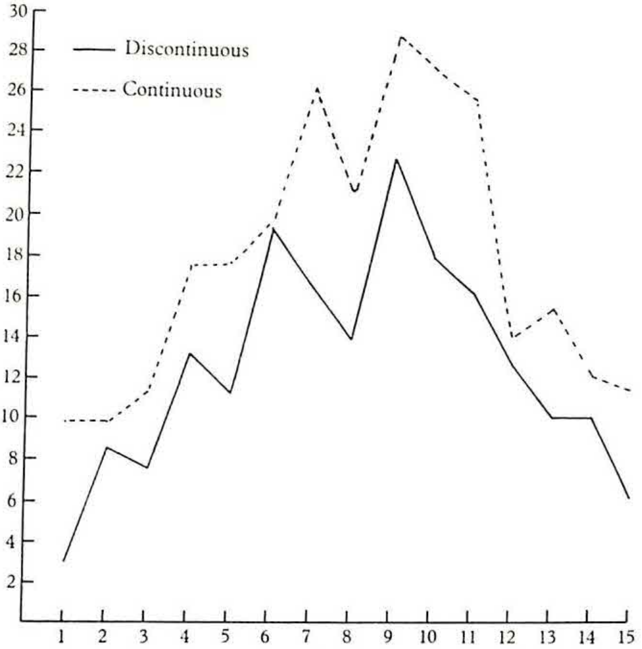


Fig. 1. Number of errors made at each syllable position until first correct anticipation with continuous and non-continuous presentation respectively.

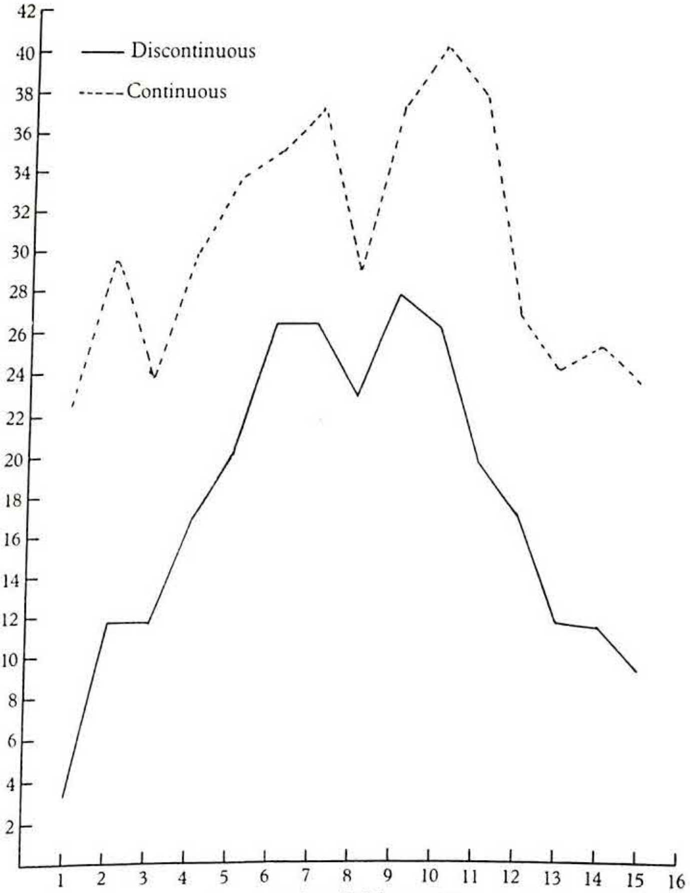


Fig. 2. Number of errors made at each syllable position during learning to mastery with continuous and non-continuous presentation respectively.

only the consolidation hypothesis, but does not seem to be in line with predictions from the Hull–Lepley hypothesis (Hull *et al.* 1940) either, as this posits specific intralist phenomena which should not apply under conditions of continuous presentation. The same may be said of the explanation in terms of competing responses (Foucault, 1928; Woodworth & Poffenberger, 1920). It also appears to contradict the Ribback & Underwood (1950) hypothesis of 'backward association'. It is possible that the correct explanation of the phenomenon may be found in connexion with the Von Restorff (1933) effect, on the assumption that the first and the last syllable in a series stand out from the rest by virtue of their position (Woodworth & Poffenberger, 1920). Even in a continuous series the subject very soon becomes aware of these position effects. The data presented do not, of course, prove this hypothesis or even support it particularly; they merely tend to go counter to all the other available hypotheses, thus leaving this as slightly less unlikely to be correct than the others.

REFERENCES

EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

FOUCALT, M. (1928). Les inhibitions internes de fixation. *Année Psychol.* **29**, 92–112.

HULL, C. L., HOWLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., and FITCH, F. B. (1940). *Mathematico-deductive Theory of Rote Learning.* New Haven: Yale University Press.

McGEOCH, J. G., & IRION, A. L. (1952). *The Psychology of Human Learning.* London: Longmans, Green and Co.

MITCHELL, M. B. (1933). The effect of serial position on the continuous memorization of numbers. *Amer. J. Psychol.* **45**, 493–4.

RESTORFF, H. v. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychol. Forsch.* **18**, 299–342.

RIBBACK, A. & UNDERWOOD, B. J. (1950). An empirical explanation of the skewness of the bowed serial position curve. *J. Exp. Psychol.* **40**, 329–35.

WOODWORTH, R. S. & POFFENBERGER, A. I. (1920). Textbook of experimental psychology, pp. 71–2. (Mimeographed typescript).

(*Manuscript received* 7 *October* 1958)

INHELDER AND PIAGET'S *THE GROWTH OF LOGICAL THINKING**

I. A PSYCHOLOGIST'S VIEWPOINT†

By JEROME S. BRUNER. *Harvard University*

In many ways, this latest in the series of books from the pen of Inhelder and Piaget is the most important yet to appear: for it represents something of a culmination of Piaget's efforts—ably aided by Professor Inhelder—to trace the intellectual operations of man from childhood to maturity. After two decades of research, Piaget and his associates have traced the development of thinking to the stage where contemporary Western man reaches his full logical powers. These powers rest in man's capacity, after a long period of biological and cultural preparation, to carry out intellectual operations that appear to be based on the implicit structures of formal logic. Somewhere between 12 and 14 years of age, with the development of ability to reflect upon thought itself, the adolescent begins to show the marks of formal thinking. He is now ready to take his place as a scientist, a thinker, a spinner of theory. If before he was able to cope with most of the problems that were immediately before him, contained so to speak in the givens of experience, now the young human being is able to cope, too, with potential problems, with issues that are not evident in the data of immediate experience, with possibilities not yet encountered. At this point, man moves beyond the powerful logic of empiricism to the infinitely more powerful logic of propositions, from a world of things present to a world of things possible. *The Growth of Logical Thinking* describes the crossing of this gap.

There is much misunderstanding of what it is that Piaget and his colleagues are trying to do—partly because their writing is outside the Anglo-American psychological tradition and partly because of a certain obscurity in Piaget's writing—attributable either to natural causes or to Piaget's wish not to 'obscure the great darkness of his subject'. Because of this misunderstanding, I should like to set forth what seems to me to be the main lines of the Geneva enterprise.

An excellent programmatic statement of the Geneva work appears in the introductory paragraphs of Piaget's slender but closely packed volume, *Logic and Psychology*. 'The aim...is to study the application of logical techniques to the psychological facts themselves, and especially to the thought structures found at different levels of intellectual development. Theoretically, it is important to ask what sort of correspondence exists between the structures described by logic and the actual thought structures studied by psychology. The question whether the structures and operations of logic correspond to anything in our actual thought, and whether the latter conforms to logical laws, is still an open one. Practically, it is important to discover in what way logic can advance psychological research....The algebra of logic can help us to specify psychological structures, and to put into calculus form those operations and structures central to our actual thought processes. Psychologists have no hesitation in using mathematics for calculating correlation coefficients, for factor analysis, etc. Now the algebra of logic is a subsystem of one of the most general fields of mathematics, that of abstract algebra. The fact that it is concerned with qualitative structures in no way detracts from its mathematical character....The psychologist for his part welcomes the qualitative character of logic, since it facilitates the analysis of the actual structures underlying intellectual operations as contrasted with the quantitative treatment of their behavioral outcome. Most "tests" of intelligence measure the latter, but our real problem is to discover the actual operational mechanisms which govern such behavior, and not simply to measure it. The algebra of logic can therefore help the psychologist, by giving him a precise method of specifying the structures which emerge in the analysis of the operational mechanisms of thought.'

Now, the task of establishing correspondences between logical structures and psychological operations of thinking is one that calls for a special approach to observation, and it is here that we find Piaget's genius. The approach consists of designing intellectual tasks that externalize thought into action in such a way that one is able to infer the logical assumptions on which the action was based and the nature of the logical system or 'structure d'ensemble' from which these assumptions were derived. This is the clinical method at its finest, clinical in its deepest sense: for the object of clinical diagnosis is to evoke for observation those forms of behaviour that indicate to us how some underlying system is operating—as a stethoscope makes observable tell-tale things about the circulatory system, and other techniques tell about other systems. What Piaget has done is to substitute the idea of logical structures for the functional systems of biology. His observational techniques are designed to permit diagnosis of these structures. In attempting to understand Piaget, one must not forget for a moment that his early formation was in biology, the kind of biology that lies midway between systematics and morphology where the use of the functional approach is a fixture of the principal method.

---

* B. Inhelder and J. Piaget, *The Growth of Logical Thinking from Childhood to Adolescence*. London: Routledge and Kegan Paul. 1958. Pp. xxvi + 356. 32*s*.

† The present article is the first of twin reviews of the Inhelder and Piaget book. The second review, 'A Logician's Viewpoint' by Charles Parsons, is to appear in the next issue of this *Journal*.

Another way of putting the matter is to say that Piaget is concerned with describing systematically the kinds of transformations that intervene between the input of a problem and the output of an attempted solution by a subject. To infer the nature of these transformations requires the externalizing of much behaviour in the form of attempted solutions, manipulation of problem materials, and 'justifications' of the adequacy of one's attempted solutions (i.e. attempted proofs). Whereas most psychologists who have attempted to describe such intervening transforms have used the empirically-oriented language of one or another psychological theory (e.g. Freud's 'dream work', Gestalt theory's 'principles of organization', Hull's 'habit family hierarchy'), Piaget proposes to describe them in terms of their correspondence to formal logical structures.

Because there has been so much specific misunderstanding of what Piaget has intended, it is perhaps well to give some simple illustrations of the search for correspondence between operations in thinking and logical structures before launching into the more complicated question of how thought goes from the stage of concrete operations to that of formal operations. The quest for an understanding of reversibility in logical and mental operations provides an instructive illustration. We begin with the observation that mental operations seem to have about them a certain notable reversibility of a kind that is not characteristic of simple overt behaviour. That is to say, when we learn how to combine classes—the class of red things with the class of round things, for example—we can immediately and without any additional learning invert the operation and decompose the combined class into its original elements. But having learned to write from left to right, we cannot automatically reverse the operation and retrace the course of our writing in the opposite direction. From this Piaget argues that the structure on which reversible mental operations is based must have a feature of *reversibility* and that any algebra of logic that we construct to describe that structure must similarly contain an operation of *inversion*. He argues, moreover, that if there is a feature of reversibility in the structure it must have certain consequences, consequences that are absent when reversibility is not present—as in certain early stages of mental functioning. One such consequence, one of central importance, is the possibility of invariance or conservation. That is to say, if the child recognizes that $A+A' = B$ in the sense that two subclasses constitute a class, and also recognizes that $A = B-A'$, then it must be the case that he also knows that the whole class B is conserved or invariant across the reversal of operations (composing and decomposing) that is performed on it. If the child has not yet reached the operational stage where reversibility is possible, B is destroyed as soon as it is divided into its components A and A'—the child cannot simultaneously deal with wholes and their parts. He is thus incapable of conservation, and it is found at this earlier stage that indeed classes and objects do not remain invariant across transformations and the young child is not capable of grasping the idea that the volume of a ball of clay does *not* change when it is rolled say into an elongated cylinder. And so, in specifying the logical structure in one instance, rules of inversion are assumed to be operative, in the other case not.

Now let us turn to the book at hand, *The Growth of Logical Thinking*. To put the matter briefly, it is argued by the authors that, as one goes from childhood to adolescence, one's range of operations is vastly increased by virtue of the fact that the adolescent is able to work directly with propositions rather than being limited to operating directly on experience, that he is able to use as a whole the logical system from which his operations are derived rather than being limited to operations taken one or several at a time, and that the system as a whole comes to guide not only his verifications of solutions but also his mode of attacking and analysing a problem.

Consider first the stage of concrete operation (ages 7–11). (I shall say very little in this review about the pre-operational thinking of children under 7 since Piaget's work at this age level is so well known.) 'In a general sense, by concrete operations we mean actions which are not only internalized but are also integrated with other actions to form general reversible systems. Secondly, as a result of their internalized and integrated nature, concrete operations are actions accompanied by an awareness on the part of the subject of the techniques and coordinations of his own behavior' (p. 6). Two forms of operation seem to be central at this stage: class inclusion and serial ordering (the former corresponding to the operation of logical inversion or negation, the latter to logical reciprocity) and both can be used with complete appropriateness to deal with events before one. Objects can be allocated to classes, and indeed classes are constructed by the 'concretely operational' child in such a way that not-yet-encountered objects and events can be placed in them. In this indirect and limited sense, the child is able to deal not only with the actual, but also with the potential. Yet the dealing is not in the abstract sense of considering the range of potential events of a full combinatorial array, but only in the sense that provisions made now are also suitable for future use. Another way of characterizing the transition from the pre-operational or 'intuitive' child to the concretely operational is that the latter comes to regard each event encountered as the result of some transformation—things are the way they are *because* of something. The position of a balance scale, for example, is seen as a resultant of weights on both ends of the balance arm rather than as a state of affairs living in its own right. It is this that gives the categories of the operational child a usefulness that goes beyond the immediately given—common factors are now seen as determining events. 'The role of *possibility* is reduced to a simple potential prolongation of the actions or operations applied to the [presently] given' (p. 249).

Now, Piaget regards the operational stages, whether concrete or formal, as being characterized by equilibrium. The reviewer is of the opinion that Piaget uses the word equilibrium in the pun sense—everything in a steady state is 'in equilibrium' in the misleading sense that everybody who is not tipsy is sober. The equilibrium of concrete operations is the 'result' of three things: (1) that events are now seen as the result of transformations (which means in a loose sense that things and events are now seen as the *result* of something having happened to make them that way); (2) that operations have now acquired completely the property of reversibility with the consequence that laws of conservation prevail with respect to the world of immediate experience if not for the world of the possible; and (3) that the structures of concrete thinking are extensible to new events, are generalizable. It seems to me that little indeed is added—save some confusing imagery—by insisting on resummarizing the definition of concrete operations and labelling this summary 'equilibrium'. It is the kind of surplus baggage that has made many students of mental life despair at trying to understand what people in Geneva mean, as in this sentence, intended by Piaget as a summary of the matter: 'The equilibrium field of concrete operations is limited in the sense that, like any equilibrium, the characteristics of such operations are determined by the compensation of potential work (operations) compatible with the system links; however, these links are limited both by the form of the operations involved and by the actual content of the notions to which they are applied' (p. 249). The French text brings this reviewer no closer to an *esprit de clarté*. I can see nothing served either in the design of experiments or in the thrust of the general theory that is contributed by the idea of equilibrium—save perhaps that it gives Piaget a comforting sense of continuity with his early biological apprenticeship. After a prolonged struggle over this point, I have come to the conclusion that what Piaget wishes to say when he writes 'equilibrium' is that each stage of operational thinking develops its own internal consistency, its own compensatory reversibility, and predictability by virtue of being based on a set of rules of operation and that these rules can be described as logical structures.

One last feature of concrete thinking before we pass on to consider the nature of formal thinking and what it accomplishes for the growing human being. It is the inability of the concrete thinker to integrate his operations into a whole by the use of higher order operations—in this sense, logical operations performed not on experience, but upon other operations. There is a hugely important point here both in the technical sense and in the more usual sense in which it has always been recognized that much is served by reflecting on one's own thought with a view toward rendering it more consistent and generic. Piaget (and I speak of him now without Professor Inhelder, for it is to his contribution that my present remarks are relevant) points out that reflexion in this sense presupposes a set of logical operations or structures that have as their particular power that they can render or translate lower order concrete transformations such as disjunction ($a \vee b$) or implication ($a \supset b$) into a *lingua franca*, of p's and q's referring to propositions, that permits the integration of operations into a structured system. The *lingua franca* corresponds, according to Piaget, to structures of propositional calculus. I am in no position to judge the depth and delicacy of Professor Piaget's representation of this change in terms of formal logic. The reader will find some severe criticisms of Piaget's representation of the formal thinking of adolescents in stage III in Mr Parsons's review—particularly concerning his ambiguous uses of truth functions and quantification theory and the looseness of his employment of the sixteen binary transformations. It is Piaget's intention, however, to characterize the shift from concrete to formal operations as a shift to a logical structure of the form of the propositional calculus. In short the adolescent learns rules about propositions and not just about things.

Consider now the consequences of the leap ahead into formal operations. First of all, it is difficult to tell from any one statement made by a subject whether he is operating concretely or formally. Take, for example, an experiment in which the subject has the task of working with rods attached to the edge of a basin of water and to the free end of each of which can be attached weights of different magnitude. The rods differ in length, in thickness, in composition (brass, steel, etc.), and in shape (round, square, rectangular). The subject has the task of determining whether or not a given rod is flexible enough to bend under a given weight and reach the water in the basin and of explaining why it does or does not do so. The subject says, for example, 'This rod bends because it is made of steel'. Can we be sure whether he is using strict implication or whether the statement merely involves class inclusion: 'This bar is steel', and 'All steel bars are flexible' in contrast to 'If a bar is steel, then it is flexible'. Clearly one cannot do this unambiguously. Rather, one looks at the whole range of propositions that a subject makes and tries to infer from them whether he is using a system of classifications or using, instead, the hypothetico-deductive method. But the best method to employ is to analyse the nature of the proofs that he employs for justifying his conclusions. If they do not go beyond observation of empirical correspondence between a class and empirical objects placed in it, then we may infer that he is operating concretely. If on the other hand, he uses a proof that employs the construction of combinations that are not represented in the objects before him, or if he constructs a relation that is not given in the properties of the objects (like a ratio), then we may assume that he is operating formally or propositionally.

The major part of the book, experiments designed principally by Mlle Inhelder, is devoted to showing what the gain is when the child shifts over to formal operations, the manner in which his proofs begin to reflect an underlying structure that is statable in terms of the sixteen binary operations of formal logic—affirmation, negation, conjunction, incompatibility, disjunction, conjunctive negation, implication, non-implication, reciprocal implication, etc. The experiments are enormously ingenious and very much in the tradition of the work that Piaget has been doing for the past decade. Since the work rests almost entirely on the acceptability of these observational experiments as proofs not only of the growth of logical thought but also of the appositeness of the logical structures that Piaget and Inhelder have claimed as corresponding to mental structures, it is worth having a close look at several of them.

In the experiment that opens the book, the child has the task of shooting a ball from a plunger against a buffer and hitting a target on the bounce. The task is to find the law that governs the relation between the angles of incidence and reflexion. To put the matter over-simply, the concrete child stops his search when he finds empirical correspondence between the angle at which he shoots the ball at the buffer and the angle at which it bounces off. After about age 11, the child begins looking for some necessity in the relationship between the two angles. Often he does this by the proof that if you point the plunger directly at the buffer wall which is at right angles to it, then it will return directly on itself—proving that one angle necessarily implies the other, that it 'must' be so. The child seems, in short, to be going in this case through an operation that is more complicated than the simpler proof of implication $p \supset q$ where one shows that $p . \bar{q}$ never occurs. Indeed Piaget and Inhelder (p. 18) suggest that the mental structure corresponds to something like:

$$p \overset{\supset}{\subset} q \text{ because } (p . q) \vee (\bar{p} . \bar{q}) \text{ are true,}$$

$$\text{and } (p . \bar{q}) \vee (\bar{p} . q) \text{ are false,}$$

where p and q state corresponding inclinations having the respective values X and Y. And when the child shoots the ball directly at the wall and it bounces back at the plunger, then:

$$(X = 0) \overset{\supset}{\subset} (Y = 0),$$

$$(X = \alpha) \overset{\supset}{\subset} (Y = \alpha),$$

where α is a determinate inclination greater than 0

$$X \overset{\supset}{\subset} Y,$$

and

$$\frown X \overset{\supset}{\subset} \frown Y,$$

where $\frown X$ is the angle of incidence and $\frown Y$ is the angle of incidence.

Thus far it is a bit like number magic, for we see little of consequence in terms of the subjects' behaviour whether they are using the logic of classes and relations or the logic of propositions as the basis of their operations. Ensuing experiments examine what some of the consequences are. A next experiment, involving bodies of various densities and sizes that either float or do not, faces the subject with the task of explaining the reason for floating or sinking. We see in this experiment that arrival at the formal stage permits subjects to eliminate contradictions from their explanations by reference to an integrated system of variables. The concretely operational child attempts to handle his explanation by a double-entry classification system of the bodies he deals with: large-heavy, small-heavy, large-light, and small-light, and discovers that small-heavy objects are the non-floating class. He does not discover the concept of density, for 'to relate the weight of the body under consideration to the weight of an equal volume of water is to invent a situation that has no empirical correlate, because only the total volume of the water in the receptacle is actually observed, whereas the conceptualization of a volume of water equal to that of the object to be compared is the product of a subtler separation of variables which once more requires hypothetico-deductive thought' (p. 37). In place of using double-entry classification, as in the concrete operations, the subject must think in terms of a logical structure that utilizes two formal operations—reciprocal implication and the notion of independence between two variables. The mental structure conforms, according to the authors, to a formal proposition based on the assumption of 'all other things being equal':

$$p.q.r \vee \bar{p}.q.\bar{r},$$

where p is the assertion that a given object floats, $\bar{p}$ the assertion that it does not, q the assertion that its volume is equal to a certain quantity of water, r the assertion that it is lighter than that body of water, $\bar{r}$ the assertion that it is heavier. The adolescent is working with operations on propositions or assertions in the form that can be stated in ordinary English as, 'A given object floats by virtue of its weight being less than that of a corresponding volume of water, or it does not float by virtue of the fact that its weight exceeds that of a corresponding volume of water'. Now what is interesting about this stage is not only that it frees the child from dealing simply with what is before him, but it also 'contains' all the earlier proofs that were useful during the stage of concrete operations. For example, the principal proof used by the concrete subject is to attempt to eliminate one of the attributes from consideration at a time (i.e. weight or size), and then to move into the elimination of the double entry classes. It is a strategy of 'negation of inclusion' which disproves that $A < B$ by showing an instance of AB'. The more formal proof presented can be shown to contain this simpler (and in the present case only partial) proof in the sense that it can also be formulated in terms of propositional logic.

Now what does all this mean psychologically? To this reviewer it suggests the diligence of Piaget's search for correspondence between logical and psychological structures. For we know that the adolescent who can operate according to the proof by relative density is also able to recognize and use the more concrete method of negation of inclusion, whereas the child who uses the latter cannot necessarily make the leap to the former. The younger child cannot get rid of the contradiction inherent in his proof, the older one can. Their respective cognitive structures must reflect this difference and Piaget proposes that it is by resort to propositional rather than data manipulation that the increase in intellectual power is brought about.

One last example will suffice, this one to show the manner in which *approach* to a problem is affected by the growth of formal thought as well as the technique of verification, as in the last example. The problem this time, again very ingenious, is to face the child with four flasks filled with colourless liquid, A, B, C and D, as well as a small bottle with a dropper, also containing colourless liquid, g. Two beakers are placed before the subject, one filled with $A+C$, the other with B. The subject does not know their content. With the subject watching, a few drops from g are placed in the two flasks. The one containing $A+C$ turns a vivid yellow, the other remains colourless. The subject is then asked to reproduce the colour, using any or all of the flasks before him. At the stage of concrete operations, the typical strategy is logical multiplication: the child adds drops of g to the content of each flask—and of course with no result. He does not seem to be able to attribute colour to a combination that goes beyond this early concrete logical operation. A bit later along in concrete thinking, the child moves ahead toward the empirical testing of n-by-n combinations—that is to say, he tries out higher order combinations but not in a manner to test adequately which particular combination is uniquely responsible for the production of colour. He is not systematically capable of eliminating variables—save for flask D which contains a liquid that prevents colour from appearing and bleaches it out when added to colour. The innovation of this more advanced concrete operation is the use of more complicated logical multiplication, but the child is still not able to operate in such a way as to test the products of such multiplication so as to discover the correct combination. He still seeks explanation in terms of the contribution of individual flasks.

When the child reaches the formal stage, a notable change occurs. He now generates a proper combinatorial—elements one at a time, two at a time, three at a time, etc., and in such an order as to eliminate systematically those *combinations* that are not adequate. The young subject is now not testing the fluids before him, the authors point out, but the hypothetical and systematic set of combinations that can be composed of an array of elements. The order of testing as well as the approach to verification has changed.

Piaget attempts to describe the change in cognitive structure as before—in terms of the child operating propositionally rather than concretely and describes the change in this case by recourse to the following logical analogues of the cognitive structures involved. To test the liquids by a systematic combinatorial, it is necessary to use all of the apparatus of the sixteen binary propositions (see above) of formal logic. That is to say, what the 'formal adolescent' is doing presupposes a system of thinking that must have all of these present. He must be using a system from which he can *derive possibilities for testing* as well as an *order of testing* and in order to do so he must be using the system-as-a-whole—the so-called '*structure d'ensemble*' which is characterized in formal logic as showing group and lattice properties (a technical logical matter that need not concern us in detail here).

Other experiments are presented, much in the same spirit, to show the manner in which formal thinking makes it possible for the adolescent to get past the errors of limited concreteness by the interpropositional operations of the sixteen binaries and the INRC transformation group. The authors add to the picture in detail, but enough has now been said so that the essence of the Inhelder–Piaget approach should be clear enough for us to proceed with a more critical intent.

First, as to the general approach. Let me ask at the outset what it means to establish a correspondence between a mental operation and some formal model of that operation—be it a model stated in terms of a logical calculus as in the present volume or in terms of idealized strategies as in the work of the reviewer (cf. *A Study of Thinking*) or in terms of an 'internalized motor-skills' model as in the case of Sir Frederic Bartlett's recent book *Thinking*, or in terms of the theory of programming as in the Logic Problem Solver of Herbert Simon and his colleagues? Good correspondence between a formal model and a process—between theory and observables, for that matter—presupposes that the model will, by appropriate manipulation, yield descriptions (or predictions) of how behaviour will occur and will even suggest forms of behaviour to look for that have not yet been observed—that are merely possible. In short, the model will simulate the behavioural resultants of an input-output process, simulate not only correct moves in problem solving, but also the characteristic 'errors' that are committed and show what brought these about.

Does Piaget's formal descriptions of pre-operational (sensori-motor) thinking, thinking by concrete operations, and thinking by interpropositional or formal operations do the job? Let me show my bias from the outset by urging that it does so remarkably well, but that it could do so remarkably better—remarkably in both cases considering where we were in this field a decade or so ago. Its failure to do remarkably better derives, I think, from the relatively static, unnecessarily static, nature of the concepts to which Piaget has committed himself—particularly the rather turgid notion of equilibrium that he introduces into the system to provide a motor for the enterprise.

With respect to the fruitfulness of the model in terms of describing both adaptively successful and unsuccessful behaviour, one cannot help but be struck by the subtlety of the manner in which the 'logic of illogic' is handled in Piaget's system, and indeed by the manner in which he predicts forms of 'pre-logical' behaviour on the basis of such broad yet specific assertions that without the development of reversibility in operational thinking, the usual forms of conservation of identity, mass, energy, etc. do not hold in the child's view of the world. It is not so much in the present volume as in the earlier *The Child's Construction of Reality* that one sees the power of this idea in Piaget's description of the development of the identity concept (conservation of identity under conditions of change and disappearance), that one grasps the importance of the idea of reversibility in cognitive operations. In the present volume, the striking shift comes between the operational stage as based on the givens of experience and operations based on propositional manipulation.

One may raise the question whether in Piaget's system (and I am indebted to Charles Parsons for bringing the matter to my attention), all 'errors' are to be regarded as regressions to earlier forms of logical operation. The authors do not assert this to be the case, but they give examples that are only of this order. What of the error whereby there is a mismatch between the logical model applied to a problem and the actual structure or requirements of that problem? An example comes to mind in modern physics; the incorrect but classical application of the principle of reciprocal implication to problems of symmetry shown by Yang to be inappropriate—that space is not universally symmetrical. These are 'mismatch' errors. Piaget gives us no account of how they come about nor what is the process whereby we locate and correct errors involving the imposition of an inappropriate logical structure. He would do well to consult the Simon work on correction programmes.

Let me now raise several questions that relate to my dissatisfaction with the equilibrium concept. The main difficulty of all such self-regulatory balances is not so much that they explain too little, but rather that they explain so much as to be useless. For if it is the case that the stage of concrete operations develops an equilibrium as earlier described, how comes it that the child gets beyond this stage to that of formal operations? Indeed, what is it altogether that impels development to move in a certain direction rather than some other? Or what impels correction for inappropriateness? Piaget is obviously not satisfied to answer that it is in the nature of the beast, nor in the nature of the brain alone. He says at one point: 'But the equilibrium attained by concrete thought covers only a relatively narrow field. Consequently, the boundaries of this field remain unstable. These two features make the elaboration of formal thought necessary' (p. 249). This may be a necessary condition for movement toward formal thought or correction, but it is scarcely a sufficient one. In another place he says, in connexion with growth not being simply something in the nature of the biological beast, 'We take as the fundamental problem of adolescence the fact that the individual begins to take up adult roles. From such a standpoint, puberty cannot be considered the distinctive feature of adolescence....The essential fact [underlying the growth of formal thought] is this fundamental social transition (and not physiological growth alone)' (pp. 335–6). After remarking that there obviously must be the development of brain structures that are in some manner isomorphic with the emergence of formal thought, the authors then go on to remark that primitive societies do not seem capable of utilizing formal structures and that the stage of formal thinking probably never emerges in their children, as likely it never did among Greek children of the classical period: 'Thus the age of 11–12 years may be, beyond the neurological factors, a product of progressive acceleration of individual development under the influence of education, and perhaps nothing stands in the way of a further reduction of the average age in a more or less distant future' (p. 337).

Piaget then asks himself what it means to fill adult roles, and in his attempt to answer his own question we see how little avail is the notion of equilibrium and how prepared he should be to adopt a less passive approach to the matter of growth. 'As opposed to the child who feels inferior and subordinate to the adult, the adolescent is an individual who begins to consider himself as the equal of adults and to judge them with complete reciprocity, on the same plane as himself. But to this first trait, two others are indissolubly related. The adolescent is an individual who is still growing but one who begins to think of the future—i.e. of his present or future work in society. Thus to his current activities, he adds a life program....The adolescent is an individual who in attempting to plan his present or future, also has the idea (from his point of view it is directly related to his plans) of changing this society....Thus it is impossible to fill an adult role without conflicts and whereas the child looks for resolutions of his conflicts in present-day compensations (real or imaginary), the adolescent adds to these limited compensations the more general compensation of a motivation for change, or even specific planning for change' (p. 339).

Here, it seems to me, lies the potential dynamism in Piaget's system, the respect in which it could be remarkably better than it is. For what is plain is that the adolescent differs from the child not simply in that he uses a propositional calculus that deals with possibilities rather than merely with actualities, but rather that he is *forced* to deal with possibility by the nature of the tasks that he undertakes and by the nature of the unfolding and development of his drives and the social connexions required for filling them. It is not equilibrium that keeps him back in the concrete operational stage and not a new equilibrium that brings him forward. It is the vicissitude of coping with demands—internal and external. The growing enterprise that is an adolescent is now operating on a different programme. Logical structures develop to support the new forms of commerce with the world. It is just as plainly the case that the pre-operational child, protected by parents, need not manipulate the world of objects unassisted until the pressure for independence is placed upon him, at which time concrete operations emerge. So the concretely operational child need not manipulate the world of potentiality (save on the fantasy level) until pressure is put upon him, at which point propositionalism begins to mark his thinking. And indeed, it is the case that such propositional thinking *never* emerges in any full-scale sense. It is no surprise, then, that children of intellectually under-privileged families or of manual workers tend to be less challenged in terms of a sense of possibilities and do not develop what we commonly speak of as an abstract gift.

In short, it seems to this reviewer that the 'operations' of which Piaget writes so brilliantly are subspecies of broader programmes or strategies, that they grow in terms of the changing objectives of such strategies and are responses to the vicissitudes through which the person is going in pursuit of his objectives. Several years ago I debated with Piaget before the *Federation Suisse de Psychologie* on the topic 'L'homme calme et l'homme agité', urging that, at least with respect to perception, it was necessary to take into account not only the constraining structural factors that operated in perceiving regardless of the state of the perceiver—for example, the minimization of surprise inherent in expectancies and the economical transforms for reducing the complexities of the perceptual field in a manner akin to that proposed by Attneave—but that one would commit an error against comprehensiveness unless one also took into account the goals for which people strove and how perceptual activity was used as an instrument in goal striving. I find myself now making the same general points with respect to Inhelder and Piaget's children and adolescents. They are talking about their tactics of thought without sufficient regard for the fact that tactics derive their direction from the strategies of which they are a part. These strategies are linked with broader motivational changes. I sense that some of the reluctance to take this step from tactics to strategy comes from Piaget's reluctance to make messy the beauty of the logical structures that conform to the cognitive structures that seem to underlie the child's and the adolescent's problem solving. The current state of motivational theory and research is not one to tempt the talents of one with such a formalistic passion as Piaget's. Yet it seems likely that the next big step forward in Piaget's thinking will not come until he attempts to extend his theory of tactics into a broader theory of strategy that takes into account the objectives toward which thought is forced to move in order to cope with the goal-striving necessary for the maintenance of a going life. The authors of this work suggest that, as adult roles must be filled, the young human is forced to cope with futurity and possibility and that the apparatus of thought responds to these pressures. We need a more elaborate account of how this comes about if we are to understand the processes that produce neurotic defences in thinking, that bring about learning blocks, that produce what can best be called 'incorrigible strategies'—failure to take any advantage of available information that happens not to conform to the specific strategy on which one is embarked.

Let me make it clear that what is said above is not in the spirit of criticism of what Piaget has already done, but rather in the form of a suggestion of what, in my opinion, he must do before his approach to cognition is complete.

One final point has to do with a matter that Piaget's teacher Claparède used to refer to as *prise de conscience*, the rendering conscious and manipulable of assumptions in terms of which one works. I was reminded of the matter by a remark made to me by one of my research students upon completing the volume now under review: 'What puzzles me is why any good 14-year-old can't get right up and give Phil 140' (Professor Quine's course in logic). There is an interesting problem here that points to a later stage in thinking, a stage beyond formal operations. It is the stage that differentiates between M. Jourdain who has learned he has spoken prose all his life and M. le professeur Jourdain armed with a theory of prose, between the polyglot and the linguistic scientist, between Piaget's 14-year-old subject and Piaget himself representing logical thinking in terms of the propositional calculus. Most adults do not progress much beyond the formal stage at which Piaget leaves his subjects in this challenging book. But those who do, make science and myth and magic. I end this altogether too lengthy review with the hope that Piaget will one day give us an account of the mysterious form of reflexivity that carries the adult from being what, in America, is called 'flat-headed' to being 'egg-headed', from being intelligent to being intelligent about intelligence. And finally, while one is expressing hopes for M. Piaget's future enterprises, it would be not amiss to hope for a volume that carries the child into adolescence along the route traversed in *Play, Dreams and Imitation* so that we may see the maturing of metaphoric thought as we have seen the maturing of logical thought in this volume.

PUBLICATIONS RECENTLY RECEIVED

Verbal Behaviour. By B. F. Skinner. London: Methuen. 1958. Pp. viii + 478. 42*s.*

This book is a description of the linguistic activity of human beings, in terms drawn from animal conditioning experiments. It does not report any appreciable amount of experimental work, but relies on everyday experience for its data. It represents, in fact, as the author himself says, a way of talking about verbal behaviour. An outline of the approach is as follows.

In animal learning situations, it is possible to make a given response depend upon the presence of a particular drive. Thus if pressing one bar produces food while pressing another bar produces water, the animal will press the first bar only when hungry, and the second when thirsty. It is also possible by suitable methods of training to make response depend upon the stimulus situations, but to make its appearance in a particular situation independent of the particular drive which is operating. Thus we might make the pressing of a certain bar produce food when the animal is hungry, water when it is thirsty, an opportunity to escape if it is being shocked, or a variety of other rewards, but do so only when the red light is on. Such a method of training will secure a high rate of bar pressing from the animal when the light is on, but not when it is off. Although the principles of behaviour may be regarded as the same in these two learning situations, in the former case the behaviour is controlled by the animal's drive and is independent of the situation: while in the latter case it is controlled by the situation and is relatively independent of the drive. Skinner regards the saying of words by human beings as behaviour which is acquired following these same principles. He therefore distinguishes in the first place a class of statements which appear when the person concerned wants something. Examples of such statements are questions, commands, requests, and so on. All statements of this type are called by Skinner *mands.* A second class of utterances are controlled by the stimulus situation, and so will appear only in particular circumstances. An example of this type of utterance is the calling of an object by its name, but a remark like 'Thank you' is equally regarded as being of this type. Instances of this second class of verbal behaviour are termed by Skinner *tacts.* These two categories are the most important ones he uses, although he makes a number of minor distinctions amongst types of verbal behaviour. Both types of response are noted as occurring frequently under the control of verbal stimulation: thus they may echo sounds which have just been heard, they may be responses to written texts, or they may be responses to earlier stimulation of a quite different type as when the response *four* is made to the verbal stimulus *two plus two.* This latter class of behaviour is known as *intraverbal.* A particularly important stimulus variable is considered to be the presence and type of audience; certain statements may be made in the presence of some people which will not in the presence of others, and it is also regarded by Skinner as generally true that verbal behaviour in the absence of an audience is punished unless internal. In putting forward these categories of analysis Skinner emphasizes that they do not correspond to traditional ones. Thus the same word may represent different types of behaviour at different times, and as has already been mentioned tacts may be nouns or quite different grammatical units. He feels that the traditional method of regarding units of speech as referring to or standing for events in the outer world is misleading and seriously limited; while it may bear some relation to behaviour involving tacts, it does not pay sufficient attention to mands nor do the lines of classification correspond to those which he prefers to use.

After setting out this analysis, he goes on to consider the multiple causation of some items of behaviour, taking numerous examples from poetry and from the type of slip made popular by Freud in *The Psychopathology of Everyday Life.* He then turns to considering a fresh category which he terms the *autoclitic.* This category contains behaviour which is descriptive of verbal behaviour itself. Some such descriptions of verbal behaviour are of course simple tacts, but there is a special class of self-descriptive behaviour which is connected with actions on the part of the listener. Thus if one prefixes the remark '*It is going to rain*' with the autoclitic '*I hear that...*' the listener is trained to respond to the second part of the sentence in a slightly different fashion from that which he would produce were the same words to be heard without a prefix. Skinner goes into the use of such autoclitics in some detail and attributes to them much of the ordering and organization of speech.

The editing which then occurs through the speaker listening to his own speech and responding with anxiety to statements which have previously been punished, is then considered. If the punishment has been so successful that the response never appears even sub vocally, there need, of course, be no anxiety, and this is described as identical with Freudian repression. Various aspects of failure of editing are discussed: logical and scientific verbal behaviour is considered under the heading of sharpened stimulus control, and finally thinking receives a chapter.

In the old behaviourist tradition sub-vocal speech is given great importance, but is not regarded as identical with thought. However, thought is identified with all behaviour, verbal or non-verbal, covert or overt. Man Thinking is simply Man Behaving. It is not quite clear why the term 'thought' is retained at all on such an analysis.

It must be admitted at this stage that the reviewer is somewhat hampered in forming a general assessment of this book by a distaste for works of this size which do not contain experiments. The merits and vices of Skinner's approach appear to be very closely similar to those of Freud's *Psychopathology of Everyday Life*, already mentioned. That is, facts must be accepted on the basis of whether they seem reasonable or not. Statements which seem to the reviewer almost certainly true are inextricably mixed with statements of fact which seem to him almost certainly false. Worst of all are the statements on the borderline, which could form excellent topics for debate. What, for example, are we to make of the following assertions, taken at random?

'Unseen audiences, as in letter writing, encourage repetitious speech' (p. 221).

'The supplementary evocation of feeble verbal responses is usually funny' (p. 286).

'Pretended slips in speech may be used as a way of making a statement while avoiding punishing consequences' (pp. 377–8).

'The style of Joseph Conrad is often difficult, and this is due to a lack of intraverbal connections from past statements to the one now being read' (p. 279).

Clearly therefore Skinner is in this book fighting psychologists of mentalistic inclinations with their own weapons. He is stating a behaviouristic attitude very much in the manner of a classical introspectionist putting forward the opposite point of view. By tackling problems of higher nervous functions he may well appeal to those who feel that behaviourists are too much concerned with muscle twitches. His book should also cause anyone who thinks of behaviourists as uncultivated and inhumane people to change his mind. The pages are liberally scattered with quotations from masters of good writing: in fact some people may feel that this tendency has been carried too far when in one place a reader can hardly get the point of an illustration unless he has read and remembers Guy de Maupassant's short story *The Necklace*.

A book of this sort must, therefore, receive an extremely personal response rather than being judged on the basis of the information it contains. So far as the reviewer is concerned, much of the analysis seems reasonable and useful. It is indeed desirable to regard utterances as instances of behaviour rather than things in themselves which are manipulated by people. The main weakness of Skinner's attitude from the point of view of a cyberneticist is the lack of emphasis on central processes within the nervous system. Throughout the book the analysis is in terms of responses controlled by stimulation, and there is a definite tone of hostility towards any suggestion that responses may be controlled by a long-lasting state of the brain, a state which is at present undetectable except by its effects. As Skinner puts it 'The programme of a radical behaviourism left no originating control inside the skin'. His hostility to the opposite view, that such control does exist, is derived from his dislike of traditional analysis of thought in terms of ideas. But any reader of contemporary journals will have noticed the increasing number of experiments in which overt responses can be explained only by means of a hypothetical mediating response within the nervous system. It is not possible to handle all the experimental data now available without recourse to such central processes. The logic of many of these experiments is similar to that of a familiar argument: the meaning of a word cannot reside in its associations, because those associations can often be seen to be different in meaning. Thus the response most strongly associated with *black* may well be *white*, and the response to *cat* may be *dog*. One should therefore suppose within the nervous system some process which occurs when one presents different words such as *cat*, *chat*, *katze*, and so on, but which is not itself an overt response to these words. Many of the experimenters on this subject refer to Osgood, and his view of mediation is of course highly relevant.

Another aspect of Skinner's emphasis on peripheral events is the underestimation of causal factors within the organism. He regards it as incorrect to speak of a person as perceiving the similarity between two situations when he uses a word from one to describe another. Rather the occurrence of the word should be ascribed to the actual external similarity between the two situations. But as we all know, not everybody can see similarities, and this is precisely why an effective use of metaphor is highly regarded. The stimulus by itself is not sufficient to produce the response and the crucial causes lie within the organism. To carry this point to its extreme, there is no reason given in the whole of Skinner's book why verbal behaviour is not shown by animals other than man. If the principles of response control given were sufficient, speech should occur in any nervous system which shows those principles, such as a rat or a pigeon. On the view that the important features of speech and thought take the form of mediating responses within the nervous system, it is by no means surprising that species differences occur and that factors within the organism control behaviour just as much as those outside.

The strength and the weakness of the book can be summed up in one of its epilogues. Skinner describes a conversation which he once had with Whitehead in which the philosopher challenged him to explain his behaviour as he sat there saying 'No black scorpion is falling upon this table'. Skinner points out that the catch in this question was that no obvious physical stimulus could produce the response 'black scorpion'. He therefore suggests that these words were indeed a response to certain recent events: since Whitehead's challenge arose out of a discussion on the subject of behaviourism. He suggests that Whitehead regarded the subject very much as most people regard scorpions, but that because of his kindliness and sympathy towards the people involved in it he felt that their hearts were in the right place and that in consequence he could deny the presence of the black scorpion. On the one hand this explanation is an extreme case of Skinner's whole analysis, since it can neither be proved nor disproved. On the other hand, even if it were true it relies implicitly on a knowledge of Professor Whitehead's own personality, and it therefore does not show that his behaviour was completely under the control of the present situation.

During the time of the Hullian ascendancy many people with an empirical approach to psychology looked on Skinner as a stalwart defender of observation as opposed to an excess of theory. This strain in his thought is still present in this book since it abounds with entertaining observations of isolated incidents from real life. Nobody can fail to retain some of these stories from a reading of the book. The reviewer particularly liked the person who did something 'in the nick of his teeth' and the recall of a well known song (at a time when the American C.I.O. was trying to organize a Farmers' Union) as 'Old Macdonald had a farm, C.I., C.I.O'. There is also the criminal who climbed through a window and clung to the sill by his fingerprints. The collection of these gems shows the respect and enjoyment of raw behaviour which have always been an honourable part of Skinner's work. But a person who has great prestige, because he has contributed so much to observed knowledge, must be resisted when he substitutes speculation about the principles of behaviour for actual experiment.

DONALD E. BROADBENT

Research Techniques in Human Engineering. By ALPHONSE CHAPANIS. Baltimore: The Johns Hopkins Press. 1959. Pp. ix + 316. $6.00.

In this volume Alphonse Chapanis has assembled together material which forms a unique contribution to the general literature, especially from the teaching standpoint, on the subject of 'Human Engineering', or—as many people in the United Kingdon would term it—'Ergonomics'. The primary aim is to inform such persons as scientists and the design engineers about human response problems, with special emphasis upon 'man-machine' systems. In his own particular way the general student of psychology will also profit from reading this book, although he may find a few technical terms beyond his immediate comprehension by virtue of their unfamiliarity. In an introductory text written with a variety of readers in mind it may be impossible to satisfy the academic psychologist by adopting a sufficiently critical approach. However, Chapanis has gone a good deal farther than most in indicating both limitations and outstanding problems without destroying the value of the book as a positive guide to current practice. The treatment and phraseology is topical, in the best sense of the word, without any signs of 'writing down'.

A text which is to form a useful review of a field previously unfamiliar to the reader necessarily has to be selective, particularly in view of the fact that the topics included are treated in sufficient detail, with concrete illustrations, to give the reader a clear grasp of the basic issues. A catalogue of the substance of the various chapters indicates the scope of this book, which includes the following:

An introductory review of the 'general philosophy of human research'—An outline of various techniques of direct observation—A review of methods of studying accidents and near-accidents—An outline of 'classical' statistical techniques including fully worked examples of multi-variate analysis—A critical survey of the nature and function of experiments in the present context—A review of the special problems of experimenting on people—A brief summary of the three main psychophysical methods and their use—An examination of 'articulation testing' (i.e. audibility of words under various conditions).

The book is the extended outcome of a good deal of co-operative thinking previously summarized in a briefer form in a monograph by Chapanis entitled *The Design and Conduct of Human Engineering Studies*, commissioned in 1954 by a committee representing the three (American) military services. The present version includes the treatment of a wide variety of examples which are *not* limited to military ones and attention is constantly directed to the general implications of the specific examples selected for discussion. Indeed, students of psychology and of engineering will find here much illustrative material not readily available elsewhere in so compact a form. For example, the chapter on 'Methods of direct observation' draws together the themes of activity sampling and its problems, process analysis, flow diagrams for man-machine systems and a selection of motion-study techniques of the traditional type. Statistical methods and problems of experimental design are outlined and discussed in a number of different places. The emphasis tends to be strictly traditional here with the topic being carried forward as far as multivariate analysis. Several excellent worked examples of Analysis of Variance of different degrees of complexity are included and the task of breaking-down and interpreting the different 'sources of variance' is discussed with lucidity. The liberal financial support provided by the sponsoring organizations has enabled the author to include a large number of highly effective diagrammatic illustrations prepared in standard form. A rather novel conception lies in the direct association of statistical significance with problems of cost accounting. However, apart from a passing reference to developments in non-parametric statistics these are largely ignored.

At times it could be argued that the author's basic philosophy of experiment appears to be too exclusively behaviouristic and operational, although this emphasis is understandable in view of the avowed aims of the book. In any case, it is clear from incidental comments that Chapanis is fully aware of the valuable insight which the more phenomenological approaches supply for the better formulation and planning of behavioural studies. A more compact statement of the methodological problems of experimental research on complex human response processes would be difficult to find. From all points of view this is a book which can be unhesitatingly recommended to all who seek information on 'Research Techniques in Human Engineering'.

ROLAND HARPER

The Psychology of Learning. By JAMES DEESE. New York, Toronto, London: McGraw-Hill. 1958 (2nd edition). Pp. vii + 367. 50*s.* 6*d.*

Theories of Learning. By ERNEST R. HILGARD. London: Methuen. 1958 (2nd edition). Pp. vii + 563. 45*s.*

These books are sufficiently well known for a review to be concerned mainly with differences between the old and the new editions.

Deese's book has lost a few pages and become more compact. A number of topics have been eliminated and the discussion of the remainder has been expanded, brought up to date, and reorganized into twelve chapters in place of seventeen. A few new topics—for example, imprinting—make their appearance and the concepts of information theory are used on occasion. A little more prominence is given to theories of learning but the empirical emphasis of the first edition remains. It is perhaps unfortunate in a book that sets out to be 'self-sufficient' that the discussion of physiological mechanisms has been reduced to a brief mention of recurrent circuits and potentiation, and that the survey of learning in different animal forms has been eliminated. Again, it is difficult to see what might justify leaving out the discussion of human learning as related to age and intelligence while retaining detailed accounts of conflict and experimental neuroses. However, the net result of revision has been to improve the book. It is better connected than it used to be and it treats selected topics more profoundly. As a result of extensive rewriting, it is even easier to read. For a balanced intelligible account of the psychology of learning at an intermediate level it can be very strongly recommended.

Hilgard's pages are larger than they were in 1948 and there are 130 more of them. The core chapters of the first edition remain, now revised, expanded and brought up to date. As theorists, Wheeler, Adams, Muenzinger, Lashley and Humphrey are no longer considered, although, of course, references to their work are scattered through the text. There is a new chapter on Freud's Psychodynamics, one on The Emergence of Mathematical Models, and two on Current Developments. The thirty pages of the chapter on Freud deal with suggestions from psycho-analysis that have a bearing on learning. They make interesting reading and there is certainly as much to be said for including Freud as Lewin, who takes up about the same space. But the result is to change the book as a whole, making it seem more of a version of 'Contemporary Schools' and less of a text on learning theories. The effect is heightened by the inclusion of more history than formerly. The chapter on mathematical models is a very valuable addition, giving a brief, clear introduction that is not too frightening. Inevitably the chapters on current developments are miscellaneous. They provide opportunities for considering the views of Sheffield, Miller, Mowrer, McClelland and others on the interpretation of reinforcement, and for introducing Spence's analysis of discrimination learning, MacCorquodale and Meehl's formalization of expectancy theory, Rotter's expectancy-reinforcement theory and Hebb's neuro-psychology. The rest of the book has been thoroughly revised. As examples of changes, the chapter on Hull deals with the 1952 postulate system and makes a surer start without Miller and Dollard's simplified theory, while the chapter on Guthrie now makes reference to the work of Voeks and Estes. As before, the final chapter is pleasantly hobby-horsical. After 450 pages of exposition, the reader rightly expects the author to open out a little.

Assuming there is no need to list the merits of a book that has been a standard text for the last ten years, it may be appropriate now to consider its future. At the end of the 1948 edition, Hilgard suggested that the passage of twenty years might make a book of this kind obsolete. Is it half-way to obsolescence? Certainly it has dated. By now the era of theory and controversy in the grand manner has passed and a good deal of the subject-matter of the book has taken on a semi-historical interest. The revised edition includes so many recent references that it will certainly keep the book alive for the next decade, but the process of bringing it up to date may not be repeatable without making it too heterogeneous and unwieldy. Many of us may hope that Hilgard will now turn his great knowledge and skill in exposition to producing some new version of *Conditioning and Learning*.

R. L. REID

Biological and Biochemical Bases of Behavior. Ed. by H. F. Harlow and C. N. Woolsey. University of Wisconsin Press. 1958. Pp. xx+476. $8.00.

This book contains the papers presented at a symposium on interdisciplinary research held at the University of Wisconsin. It is only within recent years that this custom of publishing, in book form, the material of symposia has become popular. What advantages and disadvantages stem from this form of publication? The defects are often only too readily seen in the lack of interrelationships between the different topics and the differences in the scientific and literary standards of the contributors. It is therefore pleasurable to be able to record that the peculiar merits of the symposium method are seen in the present volume. The principal merit lies in the stimulus that the book provides. It is a book into which to dip and then dip again. It is a source of information but, much more than that, it leaves the reader with an uneasy feeling. He is made aware of much interesting and indeed important work of which he should know more.

There are contributions by Harlow, Magoun, Jasper, Woolsey, Lilly, Neff and Diamond, Rose and Woolsey, Pribram, Meyer, Brady, Olds, Beach, Tower, Rosenzweig, Krech and Bennett, Sperry, Riesen, and Hebb. It is not possible to discuss all of the contributions of this strong team.

Jasper's article deals with the reticular activating systems. This is an important summary of work in this field. The anatomical basis of the recruiting system has now been somewhat clarified.

Woolsey's article deals with the organization of the motor cortex. The topological distortions of the body involved in this are fascinating. He finds that the proximal joints of the limbs are represented well anteriorly in the motor area.

Lilly records observations made on monkeys in whom the vault of the skull has been replaced by a plastic calvarium. Embedded in the shell are some 610 electrodes. It is therefore possible to stimulate any part of the convexity of the cerebral hemispheres in the conscious animal.

Neff and Diamond report their experiments on the neural basis of auditory discrimination. In doing experiments on sound localization in animals, no observer so far seems adequately to have separated the effects due to time and intensity cues respectively.

Other papers deal with non-specific brain mechanisms, the medial geniculate body, neocortical function in behaviour, psychological determinants of sparing and loss following damage to the brain, the palaeocortex and behavioural motivation, adaptive functions of palaeocortical and related structures, neural and chemical regulation of behaviour and so forth.

For the physiologically minded psychologist attention must be drawn especially to Sperry's article on 'Physiological plasticity and brain circuit theory'. After the meticulously detailed analyses in some other parts of the book it is refreshing, by way of change, to come upon some thoughtful speculations as to the concepts to be employed when thinking of brain function.

E. G. WALSH

Mental Deficiency: The Changing Outlook. Ed. by Ann M. Clarke and A. D. B. Clarke. London: Methuen. 1958. Pp. xvi+513. 45*s*.

The interest which was stimulated during the early part of the nineteenth century in the problems of mental deficiency, its nature, its causes and its cure, was mainly psychological in origin. The first institutions for defectives were schools and the training of all kinds of mentally retarded children, whether they were blind, deaf or paralysed, was the central activity. Later on, the curiosity of clinicians, paediatricians and pathologists was aroused by the strange and anomalous inhabitants of schools for defectives; in due course, in the hundred years between 1850 and 1950, a set of basic diseases which caused mental deficiency was identified, beginning with mongolism, congenital syphilis, amaurotic idiocy and tuberose sclerosis and ending with demonstrable biochemical errors as in phenylketonuria. As medical science advances we may expect that drugs, diets and surgical procedures will be brought in to alleviate or prevent the adverse physical circumstances. The problem of treatment, however, will remain in its educational, social and psychiatric aspects and it is only recently being realized how large and important is this field and how much opportunity it offers to an enthusiastic psychologist.

The manual edited by the Clarkes, with the subtitle *The Changing Outlook*, emphasizes a more hopeful and vigorous point of view than was prevalent in recent decades. The dogma that the I.Q. is a fixed innate property of an individual is no longer widely believed. Genes and the specific proteins they produce are innate, but intelligence is not one of these things and it is the product of complicated interactions of individuals with their environments. This principle pervades the book and although many hands have contributed to the work there is unity of outlook.

The series of essays which compose the book are not perfectly intergrated with respect to subject-matter. There are many points of overlapping and sometimes lack of balance. The introductory chapter by J. Tizard sketches the whole problem of mental deficiency but overemphasizes the problem of definition of psychopathy. In the second chapter, on the incidence of defect, by N. O'Connor, the balance is to some extent redressed. The facts are well selected and evaluated. As in the first chapter, the third, by Ann Clarke, is too much influenced by immediate problems connected with legislation to be of general scientific interest. In the fourth chapter, however, A. D. B. Clarke gets into his stride with a critique of intelligence testing and all the evil that goes with it.

In the next three, succeeding, chapters and in chapters IX and X, the practice of copying out of other text-books is indulged in excessively. The attempt to cover aetiology, and pathology, both generical and environmental, by secondhand methods is not convincing. Probably no harm is done in that the psychologist who reads these chapters will become aware of the existence of problems he would otherwise not have known to exist and will get a glimpse of the enormous complexity of the causes of mental defect, both social and medical. However, when the writers are dealing with first-hand experience their work is of greater value. Not everyone need agree precisely with the assessments of which tests of mental ability are the most valuable in particular situations as described by H. C. Gunzburg, but no one would doubt the genuineness of his observations and the usefulness of his comprehensive survey in chapter XI. And so it is, on the whole, with the remaining chapters where the writers evidently feel at home, as does Margaret Penwill in her account of speech disorders and their therapy.

Although this work is far from being a firm text on the psychology of mental deficiency (which should really have been its title) it is a useful and courageous contribution and should stimulate more psychologists to work in this fascinating field.

L. S. PENROSE

Mental Retardation: Its Care, Treatment and Physiological Base. By H. MAUTNER. New York and London: Pergamon. 1959. Pp. vii+280. 35*s*.

This book—or booklet, as its author modestly describes it—consists of a set of twenty-five lectures given to the staff of Pineland Hospital in the United States. The preface states that it is not a text-book, but a series of essays describing clinical pictures occurring in mentally retarded patients, and emphasizes, very properly, that the theoretical discussions are problematic, based on hypotheses and theories which may be very tentative.

The author's orientation is strongly neuro-physiological, and he summarizes much American experimental literature which has not appeared in British text-books. Thus, as a reference source for the research worker interested in the pathology of deficiency, and perhaps to a lesser extent the clinican, this volume will prove valuable. An extensive bibliography is presented, without, however, any great attempt at critical evaluation and with some surprising omissions; there are only two entries under Penrose's name, for example.

Dr Mautner is a physician, and not unnaturally his interests lie with the physical aspects of his subject. One gets the impression, however, that his philosophy is so one-sided that he ignores the social, cultural and psychological aspects not only (legitimately perhaps) as subject-matter for his book, but also in his thinking on mental deficiency. In his lecture on prophylaxis and therapy, for example,, Dr Mautner rightly mentions medical improvements such as more frequent delivery of the mother in hospital, but fails to note that, although this reduces the number of birth-injured children, it greatly increases the live birth-rate and survival of grossly deformed low-grade defectives. Benda, for example, has shown that nineteen children per thousand who as recently as 1920 would have died during their first year of life from 'congenital debility and malformations', survived in 1949. Figures such as these high-light a new and growing social problem. Moreover, only a passing and uncritical reference is made to the fact that cultural and psychological stresses may (*a*) play a part in determining some types of mental retardation, and (*b*) determine the precise manifestation of the condition in others.

Although the title of the book is wide and includes reference to care and treatment of mental retardation (reinforced by an opening quotation concerning our duty to educate and help the handicapped) there is in fact no chapter devoted to education or training, and a somewhat unjustified pessimism about psychotherapy. The vast majority of these lectures concern specific, though sometimes broad, categories or clinical conditions; such titles as 'Genetically determined diseases', 'Mongolism', 'Accidents' (particularly good), 'Embryopathy' and 'Cerebral palsy' are typical of the subject-matter of the book. It will thus be noted that little is said about half the institutionalized mentally deficient population, the high-grade, who are largely physically normal and who owe their condition to a combination of genetic and environmental factors in the main.

In summary, this is a most valuable and clearly written source-book on aspects of the pathology of mental deficiency. It will, however, be of rather little interest to the majority of psychologists, for whom Penroses's *Biology of Mental Defect* remains definitive.

A. M. CLARKE and A. D. B. CLARKE

Longitudinal Studies of Child Personality: Abstracts with Index. By ALAN A. STONE and GLORIA COCHRANE ONQUÉ. Cambridge, Mass: Harvard University Press. London: Oxford University Press. 1959. Pp. xiv+314. 40*s.*

This book consists of a brief foreword by Milton J. E. Senn, an almost equally brief introduction and about three hundred abstracts of longitudinal studies of children. These are arranged alphabetically by name of author and there is a subject index. The foreword tells us that this is the first of a series of publications planned as the outcome of an inter-disciplinary longitudinal research programme. The introduction explains that the book is primarily concerned to report studies of emotional and social development. 'Parent diaries', it adds, are not included. Almost a thousand books and articles published in or before 1955 are reported to have been read in the search for relevant material. The earliest date appears to be 1923.

Though the abstracts vary greatly in the amount of detail they offer, they tend to be longer and more informative than most of their kind. When one comes to assess the book's general usefulness, however, a doubt arises concerning the value of grouping together longitudinal studies in this way: that is, of making a methodological criterion the basis for inclusion or exclusion. The foreword makes plain the historical reason why this was done. The survey of the literature, from which this book results, was carried out as a preliminary to the planning of the main research programme and with a special view to avoiding methodological pitfalls as well as duplication of the work of others. But there is little special methodological emphasis in the published abstracts. Often, except for a reference to the time span of the research, one would hardly know that the study had been longitudinal. The abstracts concentrate on 'findings' and 'author's interpretations', and most people who consult them will no doubt be interested in precisely these things. But in this case they will probably want to know what are the findings of all research, by whatever method conducted, on the particular topic which interests them. It is unfortunate that the circumstances of this book's conception should place what one suspects will be a rather awkward limitation on its usefulness as a work of reference.

MARGARET DONALDSON

The Measurement of Value. By L. L. THURSTONE. University of Chicago Press. London: Cambridge University Press. 1959. Pp. viii+322. 56*s.* 6*d.*

This book contains twenty-seven of Professor Thurstone's papers on psychological measurement and its application in the behavioural sciences. It has two main parts, dealing with subjective measurement and attitude measurement respectively. It has been edited by Thelma Gwinn Thurstone who furnishes a preface.

A commemorative volume of this kind is usually a hodge-podge of unrelated bits of work, of very unequal value. The present volume is one of the few exceptions to this rule. Thurstone did not write much but every one of the papers included here reaches a very high standard indeed; there is no dead wood at all. He also preserved throughout his life an abiding interest in the topic which is reflected in this series of papers, the measurement of subjective qualities by rational methods. Quite apart from his later monumental work on factor analysis, his fame rests securely on such contributions as his Law of Comparative Judgement, his Phi-Gamma Hypothesis, his Measurement of Value and his Theory of Attitude Measurement. These, as well as other contributions contained in this book, are absolutely basic and should be familiar to all students of psychology; the publication of this book will make them more easily available to students in small departments without the necessary library facilities. But quite apart from any utilitarian function, the book affords a welcome opportunity of reviewing between two covers the work and contributions of one of the leading psychologists of our time who has left an indelible imprint on our ways of looking at the problems of measurement and our methods of dealing with them in our research. *Si vis monumentum, circumspice.*

H. J. EYSENCK

Social and Moral Attitudes. By S. BRACHMACHARI. Calcutta: Orient Longmans Ltd. 1954. Pp. 113. 5 rupees.

This short book is a condensed account of a Ph.D. thesis written some twenty years ago under the direction of the late J. G. Flugel at London University. It is in essence a study of attitudes, personality and upbringing, making use of questionnaires administered to about one hundred university students. Questionnaires and results are given in considerable detail and there is much of interest to be found in the actual results. Does the book justify the claim made in the foreword by Dr Flugel that it 'makes a valuable contribution to the verification of certain psychoanalytic findings by what may fairly be called "experimental" methods'? It is difficult to see how such a claim can be substantiated. Consider this quotation from the conclusions at the end of the book. 'The attitude of an individual to a certain "hero" and acceptance of the same as an ideal for the Ego are determined by the Super-Ego or by the Id. Ideal persons are selected by the Ego to compensate for the unfulfilled desire of the Id or to satisfy the Super-Ego or to rationalize the conduct of the Id.' There is no real support for these conclusions in the study, and indeed it is difficult to see how independent verification of such hypotheses could be furnished on the basis of a questionnaire study. Certainly it would be difficult to agree that this is an example of the experimental method.

More interesting than the attempt to link the findings with psychoanalysis is a study of the correlations between attitudes. Although not subjected to a factor analysis, the data nevertheless led Brachmachari to the conclusion that 'the coefficients of association in moral and social attitudes have proved the existence of a "General Progressive-Conservative Factor" in social attitudes and a "General Strict-Lax Factor" in moral attitudes'. This agrees well with much later work and anticipates, at least in part, the reviewer's own discovery of radicalism/conservatism and tough-mindedness/tender-mindedness as the two main orthogonal factors in the attitude field.

The introduction to the book gives a brief review of work in this field, but this has not been brought up to date and does not go beyond 1930; it therefore gives a quite misleading impression to the unwary reader. All in all, it would be of interest to specialists in this field, and as a historical document; to say, as the blurb does, that 'it would be of interest to the general public and will be found essential for students specially interested in psychoanalysis, social sciences and education' is a distinct overstatement.

H. J. EYSENCK

An Outline of Human Relationships. By EUSTACE CHESSER, with a Preface by CYRIL BURT. London: Heinemann. 1959. Pp. 446. 25*s*.

This is a very interesting, well-written and well-constructed popular exposition of individual, educational, social and abnormal psychology in many of their aspects, but it is predominantly oriented by consideration of their bearings on daily life and its many problems. If the book is somewhat lengthy, nevertheless its flowing and generous style of presentation gives the reader ample opportunities to assimilate the many points as they pass by while the text unfolds itself. Therefore it will appeal to the general reader of some education rather than to the student. Since it is well-informed, accurate, and broad in its approach and not biased or narrow, it can be warmly recommended.

Professor Sir Cyril Burt's distinguished and scholarly Preface is another addition to his many gifted contributions to the subject of psychology, and Dr Chesser is fortunate to have found so splendid an advocate. The Preface is well-deserved, however, and this will probably be considered the best of Dr Chesser's many works.

The book begins with a review of the biological aspects of human life and development, and passes on to the early relationships of mother and child, the development of the ego, the child in the family and at school. Adolescence and its problems are dealt with, marriage, divorce, sexuality and family planning. Dr Chesser gives some space and understanding to the psychology of the unconscious, and his grasp of it helps him through the tangles of divorce, homosexuality, prostitution, sexual anomalies and conflicts. In the later parts of his book he discusses many aspects of social psychology and religion. All in all it is an admirable guide to human problems from biological and psychological viewpoints R. W. PICKFORD

The Bristol Social Adjustment Guides. By D. H. STOTT and E. G. SYKES. London: University of London Press. 1958. (Manual—Pp. 46. 25*s.*; Specimen set of Guides—3*s.* 9*d.*; Diagnostic apparatus—9*s.* 6*d.*)

There are three Guides in this series: 'The Child in School', 'The Child in Residential Care' and 'The Child in the Family'. Each is intended to help in the diagnosis of maladjustment, and the provision of three instruments for this purpose is in keeping with Stott's view that one must start from 'bits of behaviour' as they occur in the situation in which the child is observed, not from 'traits of character', the stability and generality of which are dangerously assumed. Instead of asking the user of the Guides to indicate on a rating scale the degree, say, of 'aggressiveness', Stott offers a list of descriptive statements, grouped under various headings, and simply instructs the assessor to underline those which apply. If we take as an example 'The Child in School', we find, under the general heading 'Attitudes to other children', such statements as 'squabbles, makes insulting remarks' or 'fights viciously (bites, kicks, scratches, uses dangerous objects or weapons)'. These descriptions and others like them take the place of any inquiry about 'aggression' and in this way Stott tries to get data that are, as he puts it, 'as "raw" as possible'. His purpose in this is to avoid passing the 'bits of behaviour' through 'anyone else's classificatory or inferential sieve' before they reach his own—a purpose which is obviously sensible, even if there are limits to the possibility of achieving it. Stott argues that he can come close to this objective if his statements are confined to two classes: straightforward descriptions of actions, such as 'brings flowers, gifts frequently', and reports of signs of pleasure or distress in given situations. Basing his claim on the work of Tinbergen and Lorenz, he argues that, among the social mammals, recognition of such signs is universal and highly dependable. However, the fact that he found it necessary, when he was constructing the Guides, to change an item from 'eager to greet' to 'over-eager to greet' illustrates the difficulty of keeping out judgements of a much less elementary kind. Stott's success in reducing the influence of the particular 'classificatory sieves' of those using the Guides can, of course, be judged to some extent by the agreement obtained when different people assess the same children. There is not much evidence about this so far, but in one study carried out by the National Foundation there was definite diversity in only 5% of the comparisons made.

At the start of his work in constructing the Guides, Stott had the help of a group of house-parents in children's homes, with whom he had lengthy discussions about the behaviour of individual children. Later, teachers assisted by completing provisional versions of the Guides in respect of children they knew well, suggesting modifications of the wording and adding their own freely written descriptions of the children so that any important omissions might be detected. The critical question, of course, is introduced by this word 'important': how does one decide what is worth reporting and what is not? Stott was naturally influenced in this from the beginning by his own previous work with delinquents, but he shows himself well aware of the need to be ready to revise his notions in the course of this more general study of all varieties of maladjustment. Internal criteria play a large part, finally, in determining which items shall be retained and how they shall be grouped, but Stott insists that whatever is done must make sense in terms of the theory that is guiding the work. If some suggested change fails to do so, then either the change is not made or the theory must be amended to accommodate it: there must be no resort to blind manipulation.

It is perhaps unfortunate that the general theory of maladjustment which underlies the Guides is not expounded fully in the Manual. For a detailed statement of it, Stott refers his readers to his book *Unsettled Children and their Families*, which must be regarded as a companion volume. Briefly, the critical concept is that of the 'executive-reaction'. Stott bases his analysis of maladjustment on the notion that we possess a repertoire of instinctive reactions (sometimes he is a little more cautious and speaks of them as being 'possibly' or 'presumably' instinctive), the function of which is to make an 'unsatisfactory' situation better. Maladjustment arises when, through being brought into play too frequently, these reactions become 'over-ready' and occur inappropriately. The notion of over-readiness has therefore a very important role. Unfortunately, it gives rise to certain difficulties, both on the level of pure theory and when one considers the application of the theory to the construction of the Guides, which is, of course, our present main concern. And here the change from 'eager' to 'over-eager', which was mentioned earlier, has further illustrative value. If over-readiness is crucial to the definition of maladjustment, then should not the assessor be asked—throughout the Guides and not just in an occasional item—to judge the appropriateness of the behaviour? It is an essential part of Stott's argument that the executive-reactions are in themselves normal and quite well adjusted behaviour. Only inappropriateness makes them maladjusted. So then, if a child responds by trying, say, to escape from a situation, it would appear that we must know, before we can assess his adjustment, whether the situation is really sufficiently distressing to justify the attempt at escape. And this would seem to call for evaluation, on the part of the assessor, of a kind which is hard to reconcile with Stott's attempt to restrict all description to the two classes where he feels people's ability to judge will be in no doubt.

Stott would probably answer this objection by referring to a warning which he gives in the Manual to the effect that the validity of the Guides depends on their being used in 'a situation in which stable children are happy and thriving'. But such a situation might be acutely distressing for some children. Would attempts to render it less unsatisfactory—the executive-reactions—then be maladjusted? Does the maladjustment consist in being distressed in situations which others do not find distressing or in responding in ways appropriate to acute distress when this is not in fact experienced? If the former, then over-readiness can be judged in relation to the norm—that is, by reference to a situation in which most children thrive—and Stott's two classes of statement may well suffice for its identification. In this case, however, one would have to make it very clear that when one spoke of response to 'unsatisfactory' situations one was making a value judgement—accepting the average as a standard—and not describing the situation from the point of view of any given individual. But then if one talks, as Stott does, of *instinctive* reactions to unsatisfactory situations, must not these be unsatisfactory—or otherwise—from the point of view of the responding organism?

It would be inappropriate to pursue these perplexities further here. They seem to indicate that there is room for some development in Stott's theories and, by his own insistence, one would expect this to imply modification of his practice. However, the fact that further development seems desirable does not mean that what has been achieved is insignificant. The Guides have evidently been constructed with much care and skill, they are based on wide and thoughtful experience and, so far as the reviewer can judge in advance of the attempt to use them, they should be valuable instruments of diagnosis.

MARGARET DONALDSON

The O-Structure: An Introduction to Psychophysical Cosmology. By C. C. L. GREGORY and ANITA KOHSEN. Church Crookham: Institute for the Study of Mental Images. 1959. Pp. vi+141. 21*s*.

The ultimate aim of the writers of this book is to construct a hypothetical world-model which will spread across the holes and gaps between existing branches of knowledge and provide a single self-consistent scheme into which all the phenomena studied by the different sciences—psychological as well as physical—can be satisfactorily fitted. The book consists of three distinct parts. In the first they review the five main sectors into which modern specialization has subdivided human knowledge—those of astronomy and abstract physics, of technology and applied physics, of biology and natural history, of chance and probability, and of intuitive perception as in art, ethics, and mysticism. In the second they describe in technical detail the structure of the model they propose. The third part consists of copious notes and references bringing together a surprising array of facts and generalizations in order to illustrate or support the arguments expounded in the text.

In setting forth their main proposals the authors begin by outlining the type of scientific methodology they intend to use. In its main essentials it seems to conform with the conclusions reached by psychological writers in the recent debates on this problem (see, for example, the various papers on 'Scientific Method in Psychology' in the *British Journal of Statistical Psychology*, vol. XI, parts i and ii). A genuine science, they maintain, is to be distinguished from a pseudo-science, not in virtue of its beliefs—the assumptions which it makes or the conclusions which it upholds—but solely in virtue of the methods it adopts. The traditional conception of induction as the essential technique of an empirical science, together with the postulates which commonly go with that technique (notably the postulate of mechanical causation) they firmly reject. Instead they advocate a mode of inference, based not on the relation of subject-and-predicate, but on that of similarity (i.e. instead of 'A is B' they suggest 'A is similar to B in respect of X to the degree of Y'); and in place of a world consisting of disparate substances, such as material bodies and individual minds, they propose to substitute a relatively homogeneous universe described primarily in terms of 'information'. But, as they rightly contend, the most the scientist can hope to discover is not (as is so often assumed) the 'true nature of reality', but a more or less simplified model of reality: 'The world can only be known scientifically to the extent that its characteristics are simulated by an artificial world-model.'

After reviewing in turn the available world-models already put forward by various cosmologists, they finally select as their starting-point the 'uranoid' described by Eddington in his posthumous volume on *Fundamental Theory* (1946). This is a highly technical concept, deduced from epistemological rather than from physical principles, requiring a space-time of five dimensions, and elaborated with the aid of a novel mathematical technique suggested by Dirac's work on quantum physics. Instead of elementary particles moved by force, however, they choose for their basic concept (as Russell did in propounding his version of a 'neutral monism') the notion of an 'event'. The minimal event is described in terms suggested by Planck's quantum of 'action' (energy multiplied by duration). As the minimal duration of the atomic micro-event they take the time required for light to cross a single electron, namely 10^{-23} sec. (this they prefer to write 2^{-77} sec., since in accordance with the principles of information theory, they adopt a binary notation), the same unit, they believe, is indicated by the results arrived at by Flugel and Philpott in their investigations of mental work-curves. As the upper limit they take the time required for light to traverse the Einstein universe, i.e. about $1{\cdot}4 \times 10^9$ years. Expressed as an exponential scale the whole time-range thus covers 132 octaves. For human consciousness the smallest element of duration is what James called the 'specious present', i.e. the minimal span of attention (about $\frac{1}{16}$ sec.). The ordinary range of human perception and activity therefore falls almost exactly midway between the two extremes.

The whole universe is thus envisaged as a unique and all-embracing world-event, within which sub-events are arranged in hierarchical system. The hierarchy is organized by triads: each consists of three relata, designated A, B, and O (with affixes to indicate the level), and two relations AO and BO (hence the somewhat enigmatic title given to the 'O-structure' they propose). Thus, while mechanical events (the ideal events of particle dynamics) are thought of as obeying the 'law of least action' and radiational events as obeying the 'law of least time', so mental events are said to obey the law of 'association by similarity in hierarchical network', i.e. what might perhaps be called a 'law of least difference'.

The two concluding chapters are concerned with what the authors call the 'ostensible person', the 'potential person', and the 'morbid person'. Here, it is maintained, 'a formal, mathematical structure of personality can alone provide an adequate basis for both general and abnormal psychology; and a detailed attempt is made to apply their hypothetical scheme to elucidate the problems of perception, memory, set, instinct, and drive in general psychology, and the phenomena of conflict, dissociation, multiple personality, and the typical features of neurotic and psychotic disorders in abnormal psychology.

It will be seen that the book surveys a wide range of technical problems demanding the knowledge and skill of experienced specialists. For this the writers are excellently equipped. The senior author, formerly Director of the University of London Observatory, is a mathematician and astrophysicist; and the second author a psychologist whose researches have lain in the fields of comparative and abnormal psychology. Unfortunately, as those who have struggled through Eddington's *Fundamental Theory* are aware, the abstruse and erudite nature of the conceptions involved tax to the utmost an author's powers of lucid exposition; and, as the present writers admit, their own volume 'might have been greatly improved had its publication not been so hurried'. As a result, the reader is often doubtful whether he has rightly interpreted the inferences they wish him to draw or fully grasped their method of reasoning. Their work, however, is expressly described as an 'introduction' only; and they are themselves hoping to attempt a more complete and systematic development of their 'cosmological model' based on the axiomatic method they have laid down as the ideal. Meanwhile, their arguments well deserve the attention of the student of psychology, since they at least make clear that the old mechanistic view of the universe, which so many psychologists still suppose must form the foundation of any reputable branch of science, together with the methodological postulates on which it is based, is not only hopelessly out of date, but wholly incompatible with the facts.

CYRIL BURT

The Paradoxical Universe. By GEORGE MELHUISH. Bristol: Rankin Bros., Ltd. 1959. Pp. 160. 21*s.*

Dr Melhuish believes that the puzzles and paradoxes presented by modern physics and modern psychology amount to 'a crisis in the nature of thought itself'. His solution is to substitute a 'logic of contradiction' for a logic based on the three traditional 'laws of thought'. A somewhat similar line of escape has been discussed by Professor Quine in the United States, by Margaret Masterman in this country, and by Stéphane Lupasco in France; and indeed the problem has been familiar from the days of the Eleatic thinkers. Dr Melhuish's preference is for a methodology based on what Lupasco has called *le principe d'antagonisme.* After outlining his own development of the principle, he applies it to elucidate first the 'antinomies of the physical world' (those of time, space, and motion, and the indeterminacy-postulate of Heisenberg), secondly the peculiarities of 'mental contents' (sense-data, images, dreams, and 'the contents of the unconscious mind'), and finally the problems of 'good and evil', 'art and non-art', mortality and immortality.

Unfortunately his exposition does not possess the clarity or the logical cogency that characterize Lupasco's own discussion. Lupasco's proposal* is to discard, not only the traditional logic of substance and attribute, but also the logic of classes and relations substituted for it by the 'modern school' (i.e. that of Russell and his followers), since both are appropriate only to a 'static universe of being': instead he propounds a 'logic of process' as more appropriate to a 'dynamic universe of becoming'. His basic principle is best stated in his own words: 'un dynamisme quelconque, s'il existe, afin qu'il existe, doit impliquer un dynamisme antagoniste'. And he elaborates the consequences of his postulate, not only in verbal terms, but in a symbolic notation based on that of Russell himself.

Dr Melhuish aims at a more popular style; and contents himself with dogmatic statements and illustrative instances in place of Lupasco's explicit definitions and formal deductions. Most of his sentences have to be read, re-read, and finally translated into more precise terms, before his real meaning dawns on the reader. The obscurity is by no means due solely to the obscurity of the subject, since, when Dr Melhuish summarizes Hoyle's 'steady-state' cosmology or Russell's theory of descriptions, his fondness for paradox renders quite unintelligible what Hoyle and Russell had made perfectly clear. The careless way in which his proofs have been corrected introduces still further difficulties. The book is excellently and cheaply produced; but the errors of punctuation (e.g. in the quotation from Quine) and the frequent misspelling (e.g. San Francis for Sam Francis) at times completely distort the meaning.

Nevertheless, the psychologist who struggles through Dr Melhuish's accounts will find his arguments interesting and provocative, though most readers, I fancy, will agree that he fails to make out his case. First, many of the 'paradoxes' are by no means so novel as he supposes; Kant dealt with the antinomies of time and space; Hegel developed a dialectic of thesis and antithesis, in which contradictions were subsumed in a higher synthesis; and plausible solutions of the paradoxes of motion, infinity and continuity have been put forward which Dr Melhuish ignores. Secondly, his doctrine that 'the universe is itself *intrinsically* paradoxical' seems contradicted by the fact that, in practice, and provided the data available are adequate, predictions based on the logic of classes and of relations usually find at least an approximate fulfilment; on the other hand, it would seem quite impossible to base any verifiable predictions on the axioms which Dr Melhuish would have us accept.

CYRIL BURT

The Psychology of Thinking. By ROBERT THOMSON. Harmondsworth: Penguin. 1959. Pp. 215. 3*s.* 6*d.*

One would like to have written a lot more in reviewing this book; it is a very useful addition to the Penguin Psychology series. The author sets out 'to report the present state of knowledge in...the psychology of thinking and to give some typical samples of the data on which this particular field depends' (p. 12). The author deals particularly with reasoning or reflective thinking. Within the limits set by a book of this size the author achieves his aim with some success. The reader is left in no doubt as to the difficulty of the subject-matter and the limited achievements of psychologists working in this field; he is not allowed to lose sight of the fact that thinking is a 'polymorphous' concept (Ryle), and that thinking affects and is affected by all sorts of psychological and physiological processes.

There has been of necessity some selection of material, and no doubt some readers will wonder why this or that particular aspect of thinking was not discussed in more detail. Chapters 2–5 deal very competently with problem-solving in animals and human beings, and with conceptual processes in adults and children. The information is clearly presented and simplified without undue distortion. Chapters 6 and 7, dealing with learning operations, transfer and interference, make a less favourable impression—not because they are not well done but because the connexion between the material contained in them and the problem of thinking is less direct. The chapter on Motivation and Thought is similar in this respect. The author himself admits, p. 208, that 'psychologists have confined their attention to the simplest forms of learning...and are not in a position to interpret the more complex learning which is characteristic of thought processes'. This reviewer wonders whether, for example, chapters on Intelligence and Thinking, the Pathology of Thinking, or Personality Factors in Thinking would not have been more relevant than chapters 6 and 7. However, it is to the author's credit that these additional topics are at least mentioned, if not treated at length. Perhaps the connexion between learning and thinking would have been better presented by an account of the work of Welford and his colleagues, who studied reasoning and problem-solving processes as high-grade mental skills.

---

* *Le Principe d'Antagonisme et le Logique de l'Energie* (Paris, 1951).

Chapter 11 on Language and Thought is good if rather narrow. The last paragraph is all too brief as an indication of the extent and complexity of this field. But perhaps one is just annoyed at not finding one's favourite authors mentioned!

Two surprising omissions are (i) the lack of any reference to Spearman's account of cognition, and (ii) the lack of any reference to recent American work on creativity, e.g. by Barron, Guilford *et al.* and Anne Roe.

In general the book is sensible, up-to-date and well worth reading. This reviewer found only one printing error—a semi-humorous one on p. 146.

D. B. BROMLEY

Motivation Research. By HARRY HENRY. London: Crosby Lockwood and Son. 1958. Pp. 240. 30*s.*

With this engaging volume it is evident that motivation research has 'arrived' in Britain. This book is the first, in this country, to set out its practice, techniques and uses for the busy account executive, the market researcher, the sales director and the student of social and consumer psychology, and it has the added virtues of clarity, neatness and, at times, pungency. We are left in no doubt whatever that the wilder days of 'M.R.' with their indulgence in gross over-interpretation bolstered by Freudian shock-jargon, are largely behind us and that motivation research is growing up into something much more accurate and responsible.

This book is not directed to the expert who uses motivation-research techniques, nor is it a general text-book on consumer behaviour. Its purpose is to inform and to clarify in principle what motivation research can do and tries to do, and it is aimed at those who commission research, or receive its results. It is at its best when trying to explain, with a series of very apt illustrations, how the motivation researcher *thinks*, and how he tries to design a study that can come to grips with a problem. Motivation research is shown to be very much more than the application of personality tests to market research, and the type of problems with which it can be helpful run all the way from production policy and packaging, to marketing and advertising; to each of these broad fields a chapter is devoted, illustrating with great care the kinds of question which motivation research asks, and seeks to answer.

When the term 'motivation research' first came into use many social psychologists found that, as with 'speaking prose', they had been doing it for years without knowing it. Yet it would be foolish to dismiss M.R. as merely a set of techniques which have been in use by psychologists for many years, but which are now being applied to consumer behaviour. Fundamentally, motivation research implies a way of thinking, a novel attitude to people's economic behaviour. It starts with the willingness to probe, to find answers to 'Why'-questions without asking directly for 'reasons', and it develops, with the aid of such concepts as the 'brand image', into a systematic approach to psychological gratifications which underlie or accompany economic acts. What seems to start out as a series of truisms about human nature becomes in the end, through careful documentation and many well-chosen examples, an almost elegant exercise in rational study of the irrational. If this must give pause to the economist, it should also stimulate the psychologist who, for too long, has regarded such everyday acts as shopping, cooking, washing, smoking, drinking and eating as beneath his gaze

A. N. OPPENHEIM

The Nature of Stress Disorder. London: Hutchinson Medical Publications. 1959. Pp. xiv +298. 25*s*.

This reviewer was privileged to attend a conference on the Nature of Stress Disorder held under the auspices of the Society for Psychosomatic Research in London in May 1958. The present volume is an accurate transcript of the proceedings of this conference which, in the words of Dr Charles Newman's preface, comprised 'a less eccentric crowd than at most medical meetings'. During the meeting some valuable 'off the cuff' comments were made in the discussions following the papers and the preservation of these remarks in the book reflects the pleasant informality of the occasion.

The field of psychosomatic medicine is a patchwork affair at present, some parts of it having been tilled assiduously and others merely the object of surface scratching. This volume is appropriately divided into six distinct parts which deal with the clinical appraisal of certain stress disorders, the experimental aspects of stress, occupational factors, the family setting of the patient, genetic and constitutional factors and, finally, a lucid summing-up of the conference by Dr John Wisdom. In a number of the contributions there is an emphasis upon the experimental validation of psychosomatic hypotheses. The paper by Dr John Hambling, for instance, on the correlation of elevations of diastolic blood pressure in essential hypertension with themes of aggressive resentment is a welcome example of proof-seeking behaviour in the contemporary psychosomatic field. The experimental approach, however, brings its own problems, and questions may be raised as to whether the dual role of physiologist-psychiatrist adopted by Dr Hambling in the 'stress interviews' with his patients is in accord with the most rigorous demands of experimental science. At the clinical level there is a stimulating paper by Dr J. J. O'Dwyer on stress in industry; he discusses some of the personal problems arising from promotion, e.g. isolation from one's former colleagues, or the necessity to move to a new town with or without one's family. Dr H. N. Levitt discusses, from the general practitioner's point of view, the vital role played by domestic and neighbourhood factors in the aetiology of stress disorders. There are many other interesting contributions, too numerous to mention.

General conclusions from a heterogeneous mass of papers are propounded in masterly fashion by Dr Wisdom. Above all, he carefully assesses the value of each paper in terms of the evidence which is supplied for a particular hypothesis.

In summary, this book represents a sober attempt to portray present-day strivings within the psychosomatic field. It is attractively produced and reasonably priced.

A. BALFOUR SCLARE

Die Äusserlichkeit des Lebensstils (*ein Beitrag zur Pathologie der Ichentwicklung*). By C. Frankenstein. Amsterdam: J. M. Meulenhoff. 1959. Pp. xii+295. No price given.

It is difficult to review this book, which emanates from the Professor for the education of exceptional children at the Hebrew University, Jerusalem, because it appears to be entirely notional and does not refer at any point to experimental facts, or even the kind of observation sometimes made by psychiatrists during their work. Broadly speaking, he seems to follow Jung, and to counter the American habit of regarding extraverts as superior by denouncing them as being less highly developed than introverts. While heartily applauding this sentiment, one might wish for some evidence that this was in fact true. (The author would probably deny that this was a fair representation of his points, but it would be quite impossible to summarize the very philosophical, abstract and involved theses of his book within the narrow confines of a brief review.)

As is frequently the case in books of this type, much of what the author has to say is sensible, interesting and suggestive; many of his ideas are worth following up, particularly the stress he lays on the relationship between personality and variety of cultural factors. However, he seems to regard his hypotheses as facts rather than suggestive evaluations which might form the basis of empirical studies. In the absence of such empirical studies, the book must be judged to fall outside the field of science altogether.

H. J. EYSENCK

L. & PSY. RESE
LIBRARY